Lecture Notes on Multidisciplinary Industrial Engineering

Jiuping Xu · Mitsuo Gen
Asaf Hajiyev · Fang Lee Cooke *Editors*

Proceedings of the Eleventh International Conference on Management Science and Engineering Management

Springer

Lecture Notes on Multidisciplinary Industrial Engineering

Series editor

J Paulo Davim, Dept of Mechanical Engineering, University of Aveiro
Aveiro, Portugal

More information about this series at http://www.springer.com/series/15734

Jiuping Xu · Mitsuo Gen
Asaf Hajiyev · Fang Lee Cooke
Editors

Proceedings of the Eleventh International Conference on Management Science and Engineering Management

Springer

Editors
Jiuping Xu
Business and Administration
Sichuan University
Chengdu
China

Mitsuo Gen
Fuzzy Logic Systems Institute
Tokyo University of Science
Tokyo, Tokyo
Japan

Asaf Hajiyev
Azerbaijan National Academy of Sciences
Baku
Azerbaijan

Fang Lee Cooke
Monash University
Clayton, VIC
Australia

Lecture Notes on Multidisciplinary Industrial Engineering
ISBN 978-3-319-59279-4 ISBN 978-3-319-59280-0 (eBook)
DOI 10.1007/978-3-319-59280-0

Library of Congress Control Number: 2017942984

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Welcome to the proceedings of the Eleventh International Conference on Management Science and Engineering Management (ICMSEM 2017), which will be held from July 28 to 31, 2017, at Kanazawa, Japan. The International Conference on Management Science and Engineering Management (ICMSEM) is an annual conference organized by the International Society of Management Science and Engineering Management (ISMSEM). The conference is to foster international research collaborations in Management Science and Engineering Management as well as to provide a forum for presenting current research work in the forms of technical sessions, roundtable discussions during the conference period in a nice relaxing enjoyable atmosphere. The participants can share their academic achievements in MS and EM, and communicate with others during the conference.

The ICMSEM has been held ten times since 2007 in meeting locations across Asia, Europe, and the Americas and has had a great influence on MS and EM research. In the past ten years, the ICMSEM has been successfully held in Chengdu, Chongqing, Bangkok, Chungli, Macau, Islamabad, Philadelphia, Lisbon, Karlsruhe, and Baku. All accepted papers were put into the proceedings for each International Conference on Management Science and Engineering Management and proceedings for the last five conferences can be retrieved from EI Compendex.

This year, 939 papers from 32 countries were received and 150 papers from 20 countries accepted for presentation or poster display at the conference. These papers were from Algeria, Australia, Azerbaijan, Bangladesh, Canada, China, Indonesia, Iran, Iraq, Japan, Korea, the Republic of Moldova, Pakistan, Portugal, Spain, Thailand, the UK, the USA, Uzbekistan, and Vietnam. Each accepted paper was reviewed by three reviewers, who gave objective and helpful revision advice to the authors where necessary, which has made these conference proceedings of very high quality. The papers in the proceedings have been classified into eight sections: Computing Methodology, Data Analysis, Enterprise Operation Management, Decision Support System, Green Supply Chain, Resource Optimization Management, Risk Control, and Integrated Project Management. The key issues at the eleventh ICMSEM covered many areas of current popularity in Management Science and Engineering Management. To further encourage state-of-the-art research in the field of Management Science and Engineering Management, the ISMSEM Advancement Prize for Management Science and Engineering Management was awarded for the excellent papers at the conference.

We would like to take this opportunity to thank all participants, all of whom worked exceptionally hard to ensure this conference was a success. We want to express our sincere gratitude to the following prestigious academies and institutions for their high-quality papers and great support for the ICMSEM: the Azerbaijan Academy of Sciences, the Moldova Academy of Sciences, Academy of Sciences of the Republic of Uzbekistan, Fuzzy Logic Systems Institute, and Tokyo University of Science, Brock University. We would also like to acknowledge the assistance we received from the International Society of Management Science and Engineering Management, Kanazawa University, and Sichuan University for the conference organization. We also appreciate Springer-Verlag for the publication of the proceedings. We are grateful to Professor Yoshihiko Uesugi as the General Chair, Professor Mistuo Gen and Professor Hidetaka Nambo for their work as the Organizing Committee Chair, and Professor Takashi Oyabu as Local Arrangement Committee Chair. We appreciate the great support received from all members of the Organizing Committee, the Local Arrangement Committee, and the Program Committee as well as all participants. Finally, we would like to thank all researchers for their excellent conference papers. Because of the high quality of the research, the ISMSEM Advancement Prize for MSEM is to be awarded again for papers which have focused on innovative practical applications for Management Science and Engineering Management.

As MSEM research is in continuous development, many new MSEM development trends have emerged. Our work needs to continue to focus on MSEM development so as to encourage greater and more innovative development activity. Next year, we plan to continue the novel and successful ICMSEM and intend to increase our efforts to improve the quality of the proceedings and to recommend more excellent papers for the ISMSEM Advancement Prize. The Twelfth International Conference on Management Science and Engineering Management will be hosted by Monash University, Melbourne, Australia, in July 2018. Professor Lee has been nominated as the Organizing Committee Chair for the 2018 ICMSEM. We sincerely hope you can submit your new MSEM findings and share your ideas in Melbourne, Australia.

May 2017

Jiuping Xu
Mitsuo Gen
Asaf Hajiyev
Fang Lee Cooke

Organization

ICMSEM 2017 was organized by International Society of Management Science and Engineering Management, Sichuan University (Chengdu, China), and Kanazawa University (Kanazawa, Japan). It was held in cooperation with Lecture Notes on Multidisciplinary Industrial Engineering (LNMIE) of Springer.

Executive Committee

General Chairs

Jiuping Xu	Sichuan University, China
Yoshihiko Uesugi	Kanazawa University, Japan

Program Committee Chairs

Mitsuo Gen	Fuzzy Logic Systems Institute and Tokyo University of Science
Shavkat Ayupov	Academy of Sciences of Uzbekistan, Uzbekistan
Le Dung Trang	Abdus Salam ICTP, Vietnamese
Benjamin Lev	Drexel University, Philadelphia, USA
José Antonio de la Peña Mena	Mexican Academy of Sciences, Mexico
Asaf Hajiyev	Institute of Systems Control, Azerbaijan National Academy of Sciences, Baku, Azerbaijan
V. Cruz Machado	Universidade Nova de Lisboa, Lisbon, Portugal
Alberto Pignotti	University of Buenos Aires, Argentina
Baldev Raj	Indian National Academy of Engineering, India
Roshdi Rashed	National Center of Scientific Research, France
Haruhiko Kimura	Kanazawa University, Japan

Hayato Ohwada	Tokyo University of Science, Japan
Takashi Oyabu	Kokusai Business Gakuin College, Kanazawa, Japan

Organizing Committee Chair

Mitsuo Gen	Fuzzy Logic Systems Institute and Tokyo University of Science
Hidetaka Nambo	Kanazawa University

Local Arrangement Committee Chair

Takashi Oyabu	Kokusai Business Gakuin College, Kanazawa, Japan

Local Arrangement Committee

Ayako Sawada	Hokuriku Gakuin Junior College, Kanazawa, Japan

Program Committee

Mohammad Z. Abu-Sbeih	King Fahd University of Petroleum and Minerals, Saudi Arabia
Joseph G. Aguayo	University of Concepcion, Chile
Basem S. Attili	United Arab Emirates University, United Arab Emirates
Alain Billionnet	Ecole National Superieure Informatics for Industry and Enterprise, France
Borut Buchmeister	University of Maribor, Slovenia
Daria Bugajewska	Adam Mickiewicz University, Poland
Saibal Chattopadhyay	Indian Institute of Management, India
Edwin Cheng	Hong Kong Polytechnic University, Hong Kong
Anthony Shun Fung Chiu	De La Salle University, Philippines
Jeong-Whan Choi	Department of Mathematics, Republic of Korea
Kaharudin Dimyati	University of Malaya, Malaysia
Behloul Djilali	University of Sciences and Technology Houari Boumediene, Algeria
Eid Hassan Doha	Cairo University, Giza, Egypt
O'Regan Donal	National University of Ireland, Ireland
Siham El-Kafafi	Manukau Institute of Technology, New Zealand

Christodoulos A. Floudas	Princeton University, USA
Masao Fukushima	Kyoto University, Japan
Oleg Granichin	Sankt-Petersburg State University, Russia
Bernard Han	Western Michigan University, USA
Rene Henrion	Humboldt University, Germany
Voratas Kachitvichyanukul	Asian Institute of Technology, Thailand
Arne Løkketangen	Molde University College, Norway
Andres Medaglia	University of the Andes, Colombia
Venkat Murali	Rhodes University, South Africa
Shmuel S. Oren	University of California Berkeley, USA
Turgut Öziş	Ege University, Turkey
Panos M. Pardalos	University of Florida, USA
Gianni Di Pillo	Sapienza University of Rome, Italy
Nasrudin Abd Rahim	University of Malaya, Malaysia
Celso Ribeiro	Fluminense Federal University, Brazil
Hsin Rau	Chung Yuan Christian University, Taiwan
Jan Joachim Ruckmann	University of Birmingham, UK
Martin Skitmore	Queensland University of Technology, Australia
Frits C.R. Spieksma	Katholieke University Leuven, Belgium
Yong Tan	University of Washington, USA
Albert P.M. Wagelmans	Erasmus University Rotterdam, Netherlands
Desheng Dash Wu	University of Toronto, Canada
Hong Yan	Hong Kong Polytechnic University, Hong Kong

Secretary-General

Zhineng Hu — Sichuan University, China

Under-Secretary

Lin Zhong — Sichuan University, China

General

Zongmin Li — Sichuan University, China

Secretaries

Ning Ma, Jingqi Dai, Lurong Fan, Yan Wang

Contents

Volume 1

Data Analysis

Volume 2

Advances in Engineering Management of the Eleventh ICMSEM

Integrated Project Management

Advances in Management Science of the Eleventh ICMSEM

Advancement of Computing Methodology, Data Analysis, Enterprise Operation Management and Decision Support System Based on the Eleventh ICMSEM Proceedings

Jiuping Xu(✉)

Uncertainty Decision-Making Laboratory, Sichuan University,
Chengdu 610065, People's Republic of China
xujiuping@scu.edu.cn

Abstract. Management is playing an increasingly more important role in modern society. In particular, the development of efficient innovative managerial tools has significantly influenced social progress in management science and engineering management. In this paper, we first give a brief introduction to the eleventh ICMSEM proceedings Volume I. An analysis of the current management science topics reveals that computing methodology, data analysis, enterprise operation management, and decision support system have become key foci in the past few years. We then conduct a literature review to examine the main research in these four areas, after which the central issues in the eleventh ICMSEM Proceedings Volume I are examined using NodeXL. Finally, by analyzing the main keywords using CiteSpace, frontier management science is identified. All in all, the ICMSEM continues to provide a valuable forum for academic exchange and communication and will continue to play an important role in promoting MSEM advancements in the future.

Keywords: Computing methodology · Data analysis · Enterprise operation management · Decision support system

1 Introduction

The Eleventh International Conference on Management Science and Engineering Management (ICMSEM) in Kanazawa, Japan, has given Management Science and Engineering Management (MSEM) academics the opportunity to present their innovative findings in this increasingly popular research field. The papers in this volume demonstrate the substantial growth in interdisciplinary MSEM methodologies and practical applications. The ICMSEM aims to be the primary forum for academic researchers and practitioners to become involved in discussions on state-of-the-art MSEM research and development.

MSEM has shown tremendous growth and expansion since its beginnings. In particular, Management Science (MS) has had a long history associated with universal and impersonal management research disciplines. Scientific management,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_1

of course, owes its beginnings to the innovative ideas of Taylor and Fayol and their associates, who at the turn of the century attempted to solve industry problems through strict time monitoring, technical production methods, incentive wage payments, and rational factory organization based on efficiently structured work assignments [1,6]. In the period following the Second World War, modern analytical management methods brought the business practice into a new era. Terms such as Decision Theory, Operations Research, System Engineering, and Industrial Dynamics, which were practically unknown in the early fifties, are now as well-known as Accounting or Finance [12]. As time passed, with societal, economic, organizational, and cultural, factors playing increasingly more important roles in management practice, MS used mathematics information science, systems science, cybernetics, statistics and other theories and methods from natural science to develop innovative management and control systems. The integration of these research areas has brought significant improvements to the ICMSEM proceedings Volume I, which is mainly focused on computing methodology, data analysis, enterprise operations management, and decision support system this year.

2 Literature Review

Literature reviews give insights into the focus and new research directions in a particular research field. The widespread research attention given to MS has been mainly focused on four main areas; computing methodology, data analysis, enterprise operation management, and decision support system. Therefore, in this section, we review the pertinent research that had taken place in each of these four areas.

2.1 Computing Methodology

By providing the algorithms for management science and engineering management problems, computing methodology (CM) could be seen to be the theoretical foundation to MSEM. Through a comparative study, an appraisal of wind speed distribution prediction using soft computing methodologies has been examined for energy applications [20]. This soft computing methodologies combination meant that a more complete array from the information database could be used to develop tactical planning forecasts. Obrad Anicic et al. built an effective wind turbine noise level prediction model with the polynomial and radial basis function being applied as the kernel function for support vector regression (SVR) to estimate wind turbine noise levels [4]. Gotmare et al. presented a comprehensive review on the various swarm and evolutionary computing schemes that could be employed for system identification and digital filter design, and envisioned a quick reference for a few popular evolutionary algorithms [11]. A hybrid parallel programming model was designed for which a novel sparse matrix partitioning algorithm, a performance analysis, and an SpMV optimization on a CPU−GPU heterogeneous computing platform was developed to solve the problem [25].

Jardim and Silva presented a simple and computationally efficient methodology for computing robust dynamic equivalents that could be easily embedded in time domain simulation packages [14]. Current scholars are not only focusing on computing methodological theoretical research, but also concentrating on applications to ensure better results.

2.2 Data Analysis

Data analysis (DA) refers to the statistical analysis of large amounts of collected data to extract useful information and form conclusions. West and Post conducted research to quantify potential soil C sequestration rates for different crops with the goal of decreasing tillage intensity and enhancing rotation complexity [24]. For environmental management, carbon sequestration rates have been used in spatial modeling analyses to more accurately predict regional, national, and global C sequestration potential, and data envelopment analysis (DEA) methods have been adapted for eco-efficiency. DEA can provide substitution possibilities for different natural resources and emissions as Kuosmanen illustrated in the development of an application to assess road transportation in the three largest towns in eastern Finland [17]. To demonstrate the accuracy and potential of the methodologies, Muxika et al. used historical data, expert judgement, and multivariate analysis together with a definition of reference conditions to examine changes in the benthic communities in the Basque Country over the last decade [18]. With the development of information technology, big data has become a popular research focus because of its potential to give valued insights for enhanced decision-making. For example, big data and big data analytics methods have been used to assist decision makers better reach their targets [23]. DA is also playing an increasingly dominant role in modern society and as a result, data analysis research has yielded many competing models and contributed both conceptually and empirically to technology and organizational resource management.

2.3 Enterprise Operation Management

Enterprise operation management (EOM) generally refers to enterprise teams with the visibility and automation tools to increase returns on assets, improve management performance, and reduce enterprise costs. In the enterprise network, computing and efficient resource use can provide enterprise-wide information sharing, allowing for convenient communication between employees as well as external information access. Bi et al. investigated the impact of the emerging internet of things on enterprise systems in modern manufacturing by discussing the evolution of manufacturing system paradigms [7]. Camarinha-Matos et al. presented a discussion on the contribution of the collaborative networked organizations paradigm to the challenges faced by manufacturing systems [8]. It is undeniable that big data has attracted significant attention from researchers in the information sciences as well as policy and decision makers in governments and enterprises. Therefore, there is no doubt that business productivity and technological innovation in the future will draw on the advantages offered by big data [9,15]. Because of the promise of more automated information exchanges in networked enterprise scenarios, enterprise information systems (EIS) have become increasingly important for interoperability to increase productivity and efficiency [2]. With improvements in enterprise operations management and the application of advanced management tools, operating efficiency is being continually enhanced, thereby benefiting enterprise operations.

2.4 Decision Support System

Decision support system (DSS) are computer-based information systems that support knowledge management, decision making, and management reporting to assist managers make decisions in highly uncertain and complex environments. As is known, when making decisions, people weigh up individual opinions and combine them using a thought process to come to a final informed decision. Critically analyzing the nature and state of the history of decision support system (DSS), we can see the evolution of research and practice in areas such as personal DSS, group support systems, negotiation support systems, and knowledge management-based DSS [5]. Aiello et al. proposed a methodological approach to the development of a DSS in the specific context of Integrated Pest Management for intensive greenhouse production and provided an experimental validation based on real data [3]. A hybrid decision support system consisting of a three level humanitarian relief chain: a simulator, a rule-based inference engine, and a knowledge-based system: was designed for a humanitarian relief chain in which three main performance measures were considered for an explicit trade-off analysis between cost efficiency and responsiveness [21]. DSS can be applied to most decision making processes through the research, design and development of strategies to ensure adoption and integration.

3 The Central Issues in the Proceedings Volume I

Literature mining identifies the most pertinent scientific research in areas of particular interest [22] and has proven to be a powerful method for revealing major trends in published scientific literature across the years to allow for topic maps to be built [10]. NodeXL was designed to facilitate the learning of the concepts and methods of social network analysis using visualization as a key component [13] and is a powerful, easy-to-use interactive network visualization tool that leverages the widely available Microsoft Excel application as a platform for representing generic graphical data, performing advanced network analyses, and providing a visual exploration of the networks. NodeXL generally supports multiple social network data providers who import graphical data (node and edge lists) into an Excel sheet, and has also been widely used by researchers [16,19]. Therefore, it plays a vital role in the analysis of the keyword trends in our research. Based on the most popular research topics, we called for papers from around the world. This year, 150 papers were accepted which were divided into two proceedings volumes of 75 papers each. The significance of the keywords lies not only in the frequency or ratio but also in the key word connections that demonstrate how these papers revolve around MS concepts and metrics. Information visualization methods have also proven valuable in discovering patterns, trends, clusters, and outliers, even in complex social networks. We have now received the submissions and sorted the keywords using the software Keywords Match to perform keyword matching and to identify similarities. Finally, the processed keywords were put into the NodeXL software, and the results are shown in Fig. 1.

Fig. 1. Research topics in MS for the Eleventh ICMSEM

From this figure, it is not easy to identify the most important and most heavily researched areas. Further, from calculating and screening keyword degrees greater than six, the following graphical results were obtained. It can be seen the central issues of the eleventh ICMSEM proceedings Volume I have been mainly about computing methodology, data analysis, enterprise operations management and decision support system, as shown as Fig. 2. In the following, we focus on some related papers published in the proceedings to highlight the mainstream research this year.

There was a great deal of interest in the computing methodology section. Usman Rafi et al. discusses various scheduling terms and scheduling algorithms and proposes a new approach for scheduling based on a mixture of MQMS, a priority scheduling mechanism, and round robin scheduling. Lin et al. proposes a Bayesian network based particle swarm optimization (BNPSO) for solving a stochastic scheduling problem. Tahyudin and Nambo determine the rules of numerical association rule mining and also calculate the multi-objective function using a combination of PSO and the Cauchy distribution. Shen et al. proposes an AdaBoost algorithmic model based on a back-propagation neural network

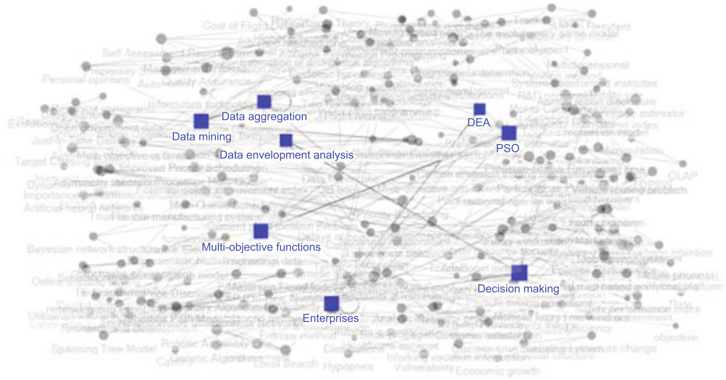


Fig. 2. Research topics in MS for the eleventh ICMSEM

for highly accurate and efficient credit scoring. Both classical and intelligence algorithms are used to develop the model solution.

The data analysis section also presents some innovative work. Yuko et al. provides a method to convert Braille books into machine-readable electronic data. Another interesting study is that in which Yamamoto researches the number of visitors in each period and their characteristics based on mobile phone user location data collected by a mobile phone company. Zeng et al. develops a severity crash count based on a hotspot identification method by extending the traditional empirical Bayes method to a generalized nonlinear model-based mixed multinomial logit approach. In addition, Hu determines the future population trends in Sichuan, China through the establishment of a population development equation. All these papers solve problems by analyzing data and developing useful tools.

For enterprise operations management, Sawada and Yoshida present a novel approach to increase the attractiveness of Facebook pages, Guo et al. uses a Helloing model to analyze the influence of pricing and profit on platform enterprises, Lu uses IPA methods to help identify the strengths and weaknesses of tourist satisfaction towards traditional culture, and Khan et al. provides cognitive support for the development of optimized employee motivation levels through consciousness raising about applied approaches and unrealistic ideas. A method that measures propensity score matching is proposed by Tang to assess how cross listing affects corporate performance. In a word, researchers have studied many aspects of enterprise operations management to improve efficiency and performance.

In the decision support system section, some practical applications to solve decision problems are presented. Pakkak proposes an integrated data envelopment analysis (DEA) and analytic hierarchy process (AHP) approach to obtain attribute weights, Sarwar et al. uses a fuzzy analytical hierarchical process (AHP) and an extent analysis method to choose an appropriate supplier under fuzzy environment, Benmessaoud et al. focuses on the monitoring of a wind farm in real time based on big data collected by the Supervisory Control and Data Acquisition (SCADA) system, and shows how the maintenance decision-making can be assured by the SCADA system, and Zhang et al. provides a corresponding pricing strategy for product crowdfunding, providing inspiration and reference for enterprises and entrepreneurs when making pricing strategy decisions. Integrating modern technology into traditional methods has become a tendency in the study of decision support system.

4 Development Trends for MS and ICMSEM

In this section, we use CiteSpace software to understand the effects of scientific research on future MS research and practice. CiteSpace is citation visualization analysis software based on the analysis of latent knowledge in scientific analysis. CiteSpace, proposed by the Chinese scholar Chen from Drexel University based on Java language, is visual document analysis software that shows the trends and tendencies within a domain of knowledge over a certain period and allows for the identification of the evolution occurring at the research frontier; it has been widely used for knowledge map visualization. The purpose of this work was to visualize and analyze the management science scientific knowledge map. In the "Web of Science" database, "management science" was used as the subject word to search and retrieve articles from 1990 to 2017, from which 62924 articles were identified. To ensure an adequate correlation for the data cleaning, TI = (Management Science) was used as the title word retrieval to exclude other databases and eliminate redundancies, thereby reducing the total to 2231 articles.

4.1 CiteSpace Analysis Results

There are 402 keyword categories and 633 edges in the map for the co-occurring keywords, in which the nodes represent each individual keyword and the lines indicate keyword co-occurrence, with the overall network density being 0.0079. To simplify the display, interactive modification was used on the visualization. By setting up "threshold = 30", we obtained a new keyword co-occurrence network map, as shown in Fig. 3.

However, on the whole, there is a loose structure in the keyword co-occurrence network as the density is not high, indicating that outside collaborations are needed to ensure that sufficient attention is paid to the research subject for international and practical needs. Deeply expanding research around a theme is vital as a lack of breadth and depth is not conducive to the sound development of management science.

In these "Co-occurring keywords", 402 categories appear. To better present the high frequency themes, the top thirty were analyzed and ranked from top to bottom in Table 1.

To investigate the use of different keywords in different time windows, the management keyword frontier research title terms were clustered into 51 categories based on the time statistics and then labelled in sequence. The first ten

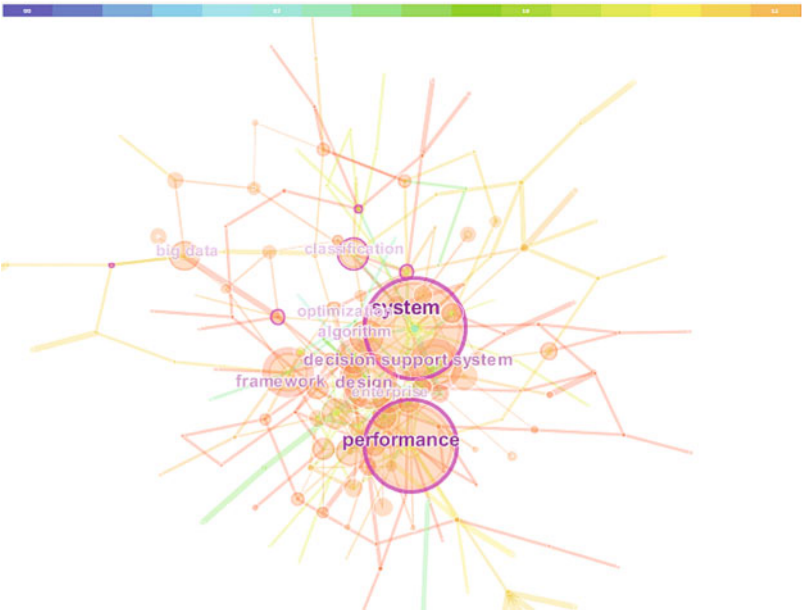


Fig. 3. Results of co-occurring keywords

classes are shown in timezone view in Fig. 4. As can be seen, management science has been in dynamic development at the points where the key word repetition is relatively high; however, there are no stable phases. There were also some indirectly related key words scattered in the small nodes around the center (not shown), which reflected the immaturity in some of the past management science research and the combinations with other research areas.

4.2 Tendency in the ICMSEM Evaluation

To further analyze the popular management science research fields, the research areas were analyzed using "Threshold = 200" in CiteSpace, the results of which are shown in Fig. 5. It can be seen that safety management, uncertainty and weighted objectives, and improved water management were the most popular study areas. From the ICMSEM papers, it can be seen that the research theme directions are similar to the research themes in journal publications. However, there are some small discrepancies between the ICMSEM proceedings volume I themes and the journal publication themes, as the proceedings are focused on innovations and more popular topics.

In addition, Table 2 indicates the highest frequency for the top thirty from the 145 categories. Of these, the most studied areas have been computer science, with the research direction having changed from computer science (2000) to environmental science & ecology (2007) to materials science & multidisciplinary research (2012) over the years. In recent years, research has focused on extending

Table 1. The first thirty co-occurring keywords categories

Frequency	Centrality	Year	Cited references
67	0.43	2002	System
55	0.31	2002	Model
46	0.04	2000	Information technology
33	0.23	2003	Performance
27	0.25	2006	Management
21	0.08	2006	Perspective
20	0.1	2001	Innovation
16	0.08	2000	Network
15	0.06	2000	Information system
15	0.04	2004	Internet
13	0.04	2006	Integration
12	0.02	2008	Impact
12	0.02	2013	Cloud computing
12	0.03	2014	Support vector regression
11	0.01	2014	Neural network
11	0.09	2007	Algorithm
11	0.06	2006	Framework
10	0.03	2008	SME
9	0	2007	Decision support system
9	0.03	2008	Organization
9	0.02	2006	Strategy
8	0.01	2014	ICT
8	0.05	2008	Prediction
8	0.02	2006	Design
7	0.03	2012	Knowledge
7	0	2006	Industry
7	0.02	2011	Architecture
7	0.04	2015	Anfi
7	0.02	2006	Technology
7	0.02	2007	Competitive advantage

the range of management science research applications in areas such as resource management, energy management, and education. Though these keywords are not in the top thirty, they are attracting increasing research attention.

When comparing Fig. 2 with the Tables 1 and 2, some of the papers in this volume not only reflect popular research areas but also introduce innovations in other areas. The articles in computing methodology (computer science theory

Fig. 4. Results of clustering keywords

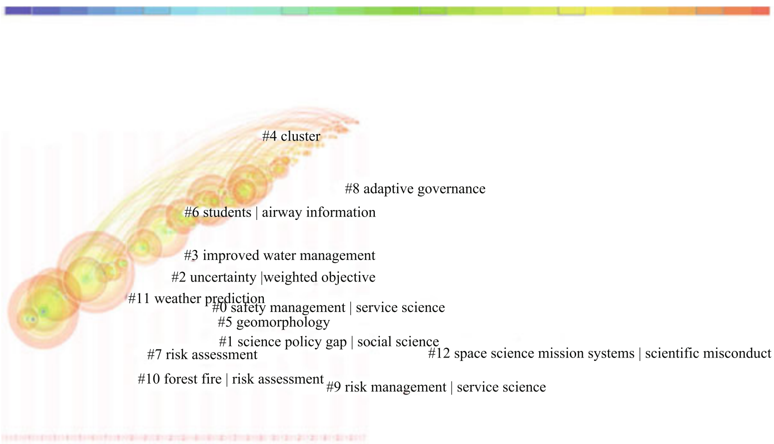


Fig. 5. The category clustering of MS timezone view

and methods, models systems, and frameworks), data analysis (statistical tools, data mining, data processes and big data), enterprise operation management (technology, information, innovation governance, business and economics) and decision support system (systems, decision making and strategies) all reflect the most pressing areas at the moment. Environmental management (ecosystems, sustainability, water resources) also has appeared as one of the main foci in the ICMSEM proceedings. When reviewing the previous years' proceedings, the

Table 2. Results of keywords cluster

Frequency	Centrality	Year	Cited references
6372	0.07	2000	Computer science
273	0.31	2000	Engineering
160	0.01	2000	Computer science, Information systems
159	0.24	2000	Engineering, Electrical & Electronic
127	0.01	2000	Computer science, Theory & Methods
117	0.03	2000	Business & Economics
114	0.33	2000	Computer science, Interdisciplinary applications
87	0.06	2001	Computer science, Artificial intelligence
86	0.1	2000	Management
60	0.09	2002	Telecommunications
51	0.02	2004	Business
49	0.01	2000	Information science & Library science
48	0.04	2001	Computer science, Software engineering
44	0.03	2001	Operations research & Management science
39	0.16	2004	Education & Educational research
32	0.05	2003	Automation & Control systems
30	0.01	2003	Computer science, Hardware & Architecture
29	0.03	2007	Environmental sciences & Ecology
28	0.13	2007	Economics
25	0.06	2000	Engineering, Industrial
22	0.1	2004	Engineering, Multidisciplinary
20	0.11	2006	Engineering, Mechanical
17	0.05	2012	Materials science
16	0	2012	Social sciences - Other topics
16	0.09	2007	Environmental sciences
16	0.04	2000	Engineering, Manufacturing
14	0.06	2012	Materials science, Multidisciplinary
14	0.05	2001	Geosciences, Multidisciplinary
14	0.07	2003	Imaging science & Photographic technology
14	0.09	2012	Social sciences, Interdisciplinary

articles included in the ICMSEM proceedings this year appear to be consistent with previous predictions. In the future, the present research vitality will hopefully result in a new round of research so that the ICMSEM continues to highlight the most popular research trends.

5 Conclusion

In this paper, we briefly introduced the four main areas covered in proceedings volume I and summarized previous research in these areas. Using the keyword analysis function in NodeXL, the most prominent topics in these four areas were identified. We also itemized the main research foci in the ICMSEM proceedings Volume I to assist readers better understand the content in this year's papers. Finally, we analyzed the MS and ICMSEM development trends using the literature analysis tool CiteSpace, which found that the research foci in the ICMSEM proceedings volume I are consistent with but slightly different from mainstream MS research. Further work is needed to identify all areas from MSEM journals so as to identify the ICMSEM expectations for the future.

Acknowledgements. The author gratefully acknowledges Jingqi Dai and Lin Zhong's efforts on the paper collection and classification, Zongmin Li and Lurong Fan's efforts on data collation and analysis, and Ning Ma and Yan Wang's efforts on the chart drawing.

References

1. Ackoff RL (1962) Scientific method: optimizing applied research decisions
2. Agostinho C, Ducq Y et al (2015) Towards a sustainable interoperability in networked enterprise information systems: trends of knowledge and model-driven technology. Comput Ind 79:64–76
3. Aiello G, Giovino I et al (2017) A decision support system based on multisensor data fusion for sustainable greenhouse management. J Cleaner Prod. doi:10.1016/j.jclepro.2017.02.197
4. Anicic O, Petkovi D, Cvetkovic S (2016) Evaluation of wind turbine noise by soft computing methodologies: a comparative study. Renew Sustain Energy Rev 56(1–2):1122–1128
5. Arnott D, Pervan G (2005) A critical analysis of decision support systems research. J Inf Technol 20(2):67–87
6. Beged-Dov AG (1967) An overview of management science and information systems. Manage Sci 13(12):817–831
7. Bi Z, Xu LD, Wang C (2014) Internet of things for enterprise systems of modern manufacturing. IEEE Trans Ind Inf 10(2):1537–1546
8. Camarinha-Matos LM, Afsarmanesh H et al (2009) Collaborative networked organizations-concepts and practice in manufacturing enterprises. Comput Ind Eng 57(1):46–60
9. Chen CLP, Zhang CY (2014) Data-intensive applications, challenges, techniques and technologies: a survey on big data. Inf Sci 275(11):314–347
10. De Bruijn B, Martin J (2002) Getting to the (c) ore of knowledge: mining biomedical literature. Int J Med Inf 67(1):7–18
11. Gotmare A, Bhattacharjee SS et al (2016) Swarm and evolutionary computing algorithms for system identification and filter design: a comprehensive review. Swarm Evol Comput 32:68–84
12. Hall AD (1962) A methodology for systems engineering. Van Nostrand, Princeton

13. Hansen D, Shneiderman B, Smith MA (2010) Analyzing social media networks with NodeXL: insights from a connected world. Morgan Kaufmann, San Francisco
14. Jardim JL, Silva AMLD (2017) A methodology for computing robust dynamic equivalents of large power systems. Electr Power Syst Res 143:513–521
15. Kambatla K, Kollias G et al (2014) Trends in big data analytics. J Parallel Distrib Comput 74(7):2561–2573
16. Kumar S, Jan JM (2012) Discovering knowledge landscapes: an epistemic analysis of business and management field in Malaysia. Procedia Soc Behav Sci 65:1027–1032
17. Kuosmanen T, Kortelainen M (2005) Measuring eco-efficiency of production with data envelopment analysis. J Ind Ecol 9(4):59–72
18. Muxika I, Angel B, Bald J (2007) Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European water framework directive. Mar Pollut Bull 55(1–6):16–29
19. Parr CS, Guralnick R et al (2012) Evolutionary informatics: unifying knowledge about the diversity of life. Trends Ecol Evol 27(2):94–103
20. Petkovi D, Shamshirband S et al (2014) An appraisal of wind speed distribution prediction by soft computing methodologies: a comparative study. Energy Convers Manage 48:133–139
21. Sahebjamnia N, Torabi SA, Mansouri SA (2016) A hybrid decision support system for managing humanitarian relief chains. Decis Support Syst 95:12–26
22. Scherf M, Epple A, Werner T (2005) The next generation of literature analysis: integration of genomic analysis into text mining. Briefings Bioinformatics 6(3):287–297
23. Sivarajah U, Kamal MM et al (2017) Critical analysis of big data challenges and analytical methods. J Bus Res 70:263–286
24. West TO, Post WM (2002) Soil organic carbon sequestration rates by tillage and crop rotation. Soil Sci Soc Am J 66(6):1930–1946
25. Yang W, Li K, Li K (2017) A hybrid computing method of SPMV on CPUCGPU heterogeneous computing systems. J Parallel Distrib Comput 104:49–60

Computing Methodology

A Comparison of Pretest, Stein-Type and Penalty Estimators in Logistic Regression Model

Orawan Reangsephet[1(✉)], Supranee Lisawadi[1], and Syed Ejaz Ahmed[2]

[1] Department of Mathematics and Statistics, Thammasat University, Bangkok, Thailand
num.stat@gmail.com

[2] Faculty of Mathematics and Science, Brock University, St. Catharines, ON, Canada

Abstract. Various estimators are proposed based on the preliminary test and Stein-type strategies to estimate the parameters in a logistic regression model when it is priori suspected that some parameters may be restricted to a subspace. Two different penalty estimators as LASSO and ridge regression are also considered. A Monte Carlo simulation experiment was conducted for different combinations, and the performance of each estimator was evaluated in terms of simulated relative efficiency. The positive-part Stein-type shrinkage estimator is recommended for use since its performance is robust regardless of the reliability of the subspace information. The proposed estimators are applied to a real dataset to appraise their performance.

Keywords: Monte Carlo simulation · Logistic regression model · Likelihood ratio test · Preliminary test estimator · Shrinkage estimator · Penalty estimator

1 Introduction

For the past few decades, simultaneous variable selection and estimation of submodel parameters has become popular. Many predictors exist to infer an interesting response in the initial model. Some of these predictors may be inactive and not influential; these should be excluded from the final model that represents a sparsity pattern in the predictor space to achieve parsimony, flexibility and reliability. Several researchers, following this information in statistical modeling, have used either the full model or a candidate submodel.

The logistic regression model also called the logit model, is the most widely used for an analysis of the independent binary response data in medical, engineering, and other studies. This model assumes that the logit of the response variable can be modelled by a linear combination of unknown parameters $x_i'\beta$

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_2

where $\mathbf{x_i'} = (x_{i1}, x_{i2}, \cdots, x_{ip})$ is a $p \times 1$ vector of the p predictors for the i^{th} subject and $\beta = (\beta_1, \beta_2, \cdots, \beta_p)'$ is a $p \times 1$ vector of regression parameters. Detailed information on logistic regression can be found in the books by Hilbe [8] and Hosmer and Lameshow [10].

In this article, we consider the problem of estimating the logistic regression model when the response variable may be related to many predictors, some of which may be inactive. Prior information about inactive predictors may be incorporated in the full model to produce the candidate submodel.

The pretest (preliminary test) estimation strategy, is inspired by Bancroft, and the shrinkage estimation strategy, is inspired by Stein, efficiently combine both full model and submodel estimators in an optimal way to achieve an improved estimator. Numerous authors have discussed the pretest, shrinkage, and penalty estimation strategies in many fields including Ahmed and Amezziane [2], Ahmed and Yüzbaşı [4], Al-Momaniet et al. [5], Gao, Ahmed, and Feng [6], Hossain, Ahmed, and Doksum [12], and Yüzbaşı and Ahmed [16,17]. For a logistic regression model, shrinkage estimators and three penalty estimators as LASSO, adaptive LASSO and SCAD were considered by Hossain and Ahmed [11] and Lisawadi, Shah, and Ahmed [13] considered the pretest estimation.

As we know, ridge regression (Hoerl and Kennard [9]) has been widely used when there are many possible predictors to achieve the precision of an estimate. Ahmed et al. [3] found that the ridge regression is highly efficient and stable when there are many predictors with small effect. Hence, we suggest the ridge regression for a logistic regression model. In this article, we propose the pretest and shrinkage estimators in the logistic regression model when it is priori suspected that parameters may be restricted to a subspace and compares the resulting estimators to the classical maximum likelihood estimator as well as the penalty estimators, i.e. LASSO estimator and ridge regression. Monte Carlo simulation study is carried out using the simulated relative efficient to appraise the performance of the proposed estimators.

To further illustrate the proposed estimators in the logistic regression model, we apply the proposed estimator to the South African heart disease data set and provide a bootstrap approach to compute simulated relative efficiency (SRE) and simulated relative prediction error (SPE) of the estimators. The detail of this data set will be described in the Sect. 4. Hossain, Ahmed, and Doksum [12] also considered this data in the generalized linear model via the pretest estimator, positive-part Stein-type shrinkage estimator, and three penalty estimators as LASSO, adaptive LASSO, and SCAD The performance of these estimators are evaluated in terms of simulated relative efficient (SRE).

Under the prior information about inactive predictors, the full parameter vector β can be partitioned as $\beta = (\beta_1', \beta_2')'$ where β_1 and β_2 represent a $p_1 \times 1$ active parameter and a $p_2 \times 1$ inactive parameter subvector, respectively, such that $p = p_1 + p_2$. Therefore, our interest lies in the estimation of the active parameter subvector β_1 when the information on β_2 is readily available. In other words, this information about the inactive parameters may be used to estimate the active parameter subvector β_1 when their values are near to some specified value β_2^0. Without the loss of generality, it is plausible that β_2 may be set to a zero vector, $\beta_2 = 0$. Keep in mind that the candidate submodel estimator is more efficient than the full model estimator when the candidate submodel is correct. On the other hand, the submodel estimator may not be reliable and become considerably inefficient when the candidate submodel incorrectly represents the data at hand.

The remainder of this article is organized as follows; the model and the efficient estimation strategies are proposed in Sect. 2, the results of a Monte Carlo simulation study are reported in Sect. 3, real data applications are described in Sect. 4, and finally, discussions and conclusions are presented in Sect. 5.

2 Model and Estimation Strategies

Let $y_1, y_2, \cdots, y_n$ be independent binary response variables which contain only two possible outcomes, and $\mathbf{x_i} = (x_{i1}, x_{i2}, \cdots, x_{ip})'$ is a $p \times 1$ predictors vector for the i^{th} subject and $i = 1, 2, \cdots, n$. The simplest idea would be to let z_i be a linear function of the predictors, suppose

$$z_i = \mathbf{x_i'}\beta, \tag{1}$$

where β is a $p \times 1$ vector of regression coefficients. Thus, the logistic regression model assume that

$$P\left(y_i = 1|\mathbf{x}_i\right) = \pi\left(z_i\right) = \frac{\exp\left(z_i\right)}{1 + \exp\left(z_i\right)} = \frac{\exp\left(\mathbf{x}_i'\beta\right)}{1 + \exp\left(\mathbf{x}_i'\beta\right)}. \tag{2}$$

The log-likelihood function of the logistic regression model is given by

$$l\left(\beta\right) = \sum_{i=1}^{n} y_i \ln\left\{\pi(\mathbf{x}_i'\beta)\right\} + \sum_{i=1}^{n}(1 - y_i)\ln\left\{1 - \pi\left(\mathbf{x}_i'\beta\right)\right\}. \tag{3}$$

The derivative of the log-likelihood function with respect to β is obtained by solving the score equation:

$$\frac{\partial l\left(\beta\right)}{\partial\beta} = \sum_{i=1}^{n}\left[y_i - \pi(\mathbf{x}_i'\beta)\right]\mathbf{x}_i = 0. \tag{4}$$

2.1 The Unrestricted and Restricted Maximum Likelihood Estimator

The unrestricted maximum likelihood estimator (UE) of the parameter vector β denoted by $\hat{\beta}^{UE}$ is obtained by solving the non-linear score Eq. 4, and this can be solved by using an iterative method like Newton-Raphson.

Under the certain regularity conditions of maximum likelihood estimator (MLE), Gourieroux and Monfort [7] showed that $\hat{\beta}^{UE}$ is a consistent estimator of β and asymptotically normally distributed with a variance-covariance matrix $(I(\beta))^{-1}$, where $I(\beta)$ is the information matrix which is defined as

$$I(\beta) = \sum_{i=1}^{n} \pi(\mathbf{x}_i^{'}\beta)\left\{1 - \pi(\mathbf{x}_i^{'}\beta)\right\}\mathbf{x}_i\mathbf{x}_i^{'}. \tag{5}$$

The restricted maximum likelihood estimator (RE) of β denoted by $\hat{\beta}^{RE}$ can be obtained by maximizing the log-likelihood function (3) under the subspace restriction $\beta_2 - \beta_2^0 = 0$.

2.2 The Linear Shrinkage Estimator

The linear shrinkage estimator (LS) of β denoted by $\hat{\beta}^{LS}$ is a linear combination of the unrestricted and restricted estimator, that is

$$\hat{\beta}^{LS} = \lambda\hat{\beta}^{RE} + (1-\lambda)\,\hat{\beta}^{UE}, \lambda \in [0,1], \tag{6}$$

where λ defines the degree of confidence in the given prior information and is a fixed constant. The linear shrinkage estimator shrinks $\hat{\beta}^{UE}$ toward $\hat{\beta}^{RE}$. If $\lambda = 0$, then LS simplifies to an unrestricted estimator, while it simplifies to a restricted estimator when $\lambda = 1$. The performance of the linear shrinkage estimator is better than the unrestricted and restricted MLE in some part of the parameter space.

2.3 The Preliminary Test Estimator

The preliminary test estimator or pretest estimator (PT) of β denoted by $\hat{\beta}^{PT}$ is defined as

$$\hat{\beta}^{PT} = \hat{\beta}^{UE} - \left(\hat{\beta}^{UE} - \hat{\beta}^{RE}\right) I\left(\mathscr{L}_n \leq \mathscr{L}_{n,\alpha}\right), \tag{7}$$

where $I(.)$ is an indicator function, and $\mathscr{L}_{n,\alpha}$ is the α-level critical value of the exact distribution of a suitable test statistic $\mathscr{L}_n$ under $H_0 : \beta_2 = \beta_2^0$. For testing $H_0 : \beta_2 = \beta_2^0$, the likelihood ratio statistic $\mathscr{L}_n$ is suggested:

$$\mathscr{L}_n = -2\log\left(\frac{L\left(\hat{\beta}^{RE}\right)}{L\left(\hat{\beta}^{UE}\right)}\right) = 2\left(l\left(\hat{\beta}^{UE}\right) - l\left(\hat{\beta}^{RE}\right)\right), \tag{8}$$

where $l(\hat{\beta}^{UE})$ and $l(\hat{\beta}^{RE})$ are values of the log-likelihood at the unrestricted and restricted estimates, respectively. Under H_0, the distribution of $\mathscr{L}_n$ converges to Chi-square distribution with p_2 degree of freedom as $n \to \infty$.

Clearly, the pretest estimator takes the value of the unrestricted estimator when the test statistic lies in a rejection region, otherwise, it takes the value of the restricted estimator. This estimator has limits due to the large size of the pretest.

2.4 The Shrinkage Pretest Estimator

The shrinkage pretest estimator (SP) of β denoted by $\hat{\beta}^{SP}$ is defined by replacing the restricted estimator with the linear shrinkage estimator in Eq. (7), that is

$$\hat{\beta}^{SP} = \hat{\beta}^{UE} - \left(\hat{\beta}^{UE} - \hat{\beta}^{LS}\right) I\left(\mathscr{L}_n \leq \mathscr{L}_{n,\alpha}\right). \tag{9}$$

An alternative form of the estimator is

$$\hat{\beta}^{SP} = \hat{\beta}^{UE} - \lambda\left(\hat{\beta}^{UE} - \hat{\beta}^{RE}\right) I\left(\mathscr{L}_n \leq \mathscr{L}_{n,\alpha}\right). \tag{10}$$

Ahmed [1] found that the shrinkage pretest estimator significantly improves upon the pretest estimator in terms of size α, and it dominates the unrestricted estimator in a large portion of the parameter space. For $\lambda = 1$, the pretest estimators are used to estimate the parameter, while we use a UE as $\lambda = 0$. Generally, the estimators based on the pretest strategy are biased and inefficient when the null hypothesis does not hold.

2.5 The Stein-Type Shrinkage Estimator

The Stein-type shrinkage estimator which combines the unrestricted and the restricted estimator in an optimal way to dominate the unrestricted estimator. The Stein-type shrinkage estimator (S) of β denoted by $\hat{\beta}^{S}$ is given as follows

$$\hat{\beta}^{S} = \hat{\beta}^{RE} + \left(1 - (p_2 - 2)\,{\mathscr{L}_n}^{-1}\right)\left(\hat{\beta}^{UE} - \hat{\beta}^{RE}\right), p_2 \geq 3, \tag{11}$$

alternatively,

$$\hat{\beta}^{S} = \hat{\beta}^{UE} - (p_2 - 2)\,{\mathscr{L}_n}^{-1}\left(\hat{\beta}^{UE} - \hat{\beta}^{RE}\right), p_2 \geq 3. \tag{12}$$

For some insight to this estimator, we refer to Hossain, Ahmed, and Doksum [12], Yüzbaşı and Ahmed [17] among others. The Stein-Type shrinkage estimator will provide uniform improvement over the unrestricted estimator. However, the Stein-type shrinkage estimator tends to over-shrink the unrestricted estimator towards the restricted estimator when the test statistic $\mathscr{L}_n$ is very small in comparison with $p_2 - 2$. To avoid the over-shrink behavior of this estimator, the truncated version is suggested which is called the positive-part Stein-type shrinkage estimator.

2.6 The Positive-Part Stein-Type Shrinkage Estimator

The positive-part Stein-type shrinkage estimator (S^{+}) of β denoted by $\hat{\beta}^{S^{+}}$ is a convex combination of the unrestricted and restricted estimator, that is

$$\hat{\beta}^{S^{+}} = \hat{\beta}^{RE} + \left(1 - (p_2 - 2)\,\mathscr{L}_n^{-1}\right)^{+}\left(\hat{\beta}^{UE} - \hat{\beta}^{RE}\right), p_2 \geq 3, \tag{13}$$

where $a^+ = \max\{0, a\}$. Alternatively, it can be written in the following conical form as

$$\hat{\beta}^{S^+} = \hat{\beta}^{RE} + \left(1 - (p_2 - 2)\,\mathscr{L}_n{}^{-1}\right)\left(\hat{\beta}^{UE} - \hat{\beta}^{RE}\right) I\left(\mathscr{L}_n < (p_2 - 2)\right), p_2 \geq 3. \quad (14)$$

The positive-part Stein-type shrinkage estimator is particularly important to control the over-shrinking inherent in $\hat{\beta}^S$.

2.7 The LASSO Estimator

Tibshirani [15] introduced the LASSO estimator of β which minimizes the negative log-likelihood in Eq. (3) under the L_1 constraint. It can be defined as

$$\hat{\beta}^{\text{LASSO}} = \operatorname{argmin}_{\beta}\left\{-l(\beta) + \gamma \sum_{i=1}^{p} |\beta_i|\right\}, \gamma \geq 0, \quad (15)$$

where λ is the tuning parameter which controls the amount of a shrinkage. The LASSO shrinks some coefficients to exactly zero. Therefore, LASSO procedure performs variable selection and parameter estimation simultaneously.

2.8 The Ridge Regression Estimator

Hoerl and Kennard [9] proposed the ridge regression estimator of β which minimizes the negative log-likelihood in Eq. (3) under the L_2 constraint. It can be defined as

$$\hat{\beta}^{RIDGE} = \arg\min_{\beta}\left\{-l(\beta) + \gamma \sum_{i=1}^{p} {\beta_i}^2\right\}, \gamma \geq 0, \quad (16)$$

where λ is a tuning parameter which controls the amount of shrinkage. The ridge regression estimator always keeps all the predictors in the model; thus, this estimator cannot produce a parsimonious model.

3 Monte-Carlo Simulation Studies

In this section, we carry out a Monte Carlo simulation to compare the performance of the pretest, Stein-type and penalty estimators in terms of the quadratic risk, namely mean squared error (MSE). Our simulations are based on a logistic regression model with the sample size $n = 250$. A binary response data is generated from the following model

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \mathbf{x_i}'\beta = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}, \quad i = 1, 2, \cdots, n, \quad (17)$$

where $p_i = P(y_i = 1|\mathbf{x_i})$ and the predictor values $\mathbf{x_i}$ have been drawn from a standardized multivariate normal distribution.

We consider the hypothesis $H_0 : \beta_2 = 0$. We partition the parameter vector β as $\beta = (\beta_1', \beta_2')'$ and where β_1 and $\boldsymbol{\beta_2}$ represent a $p_1 \times 1$ and a $p_2 \times 1$ vector, respectively, such that $p = p_1 + p_2$. We set the true value of $\beta = (\beta_1', \beta_2')' = (\beta_1', 0)'$ with $\beta_1' = (1.90, -1.05, 0.25, -0.78)$.

The value of λ is set to 0.25, 0.50, and 0.75. The value of significance level α is set to 0.01, 0.05, 0.10 and a higher value 0.35.

We now define the parameter Δ^* representing the distance between the simulated model and the candidate submodel estimator by

$$\Delta^* = \left(\beta - \beta^{(0)}\right)' \left(\beta - \beta^{(0)}\right) = \sum_{i=1}^{p} \left(\beta_i - \beta_i^{(0)}\right)^2, \tag{18}$$

where $\beta^{(0)} = \left(\beta_1^{'}, 0\right)^{'}$ and β is the true parameter in the simulated model. Samples were generated using Δ^* between 0 and 4.

The number of replications in the simulation was initially varied and it was determined that $N = 1,000$ iterations were adequate to obtain a stable result for each combination of parameters.

Based on the simulated data, we estimated the MSE of all the proposed estimators. The performance of the estimators was evaluated using the notion of simulated relative efficient (SRE), which is the MSE relative to the MSE of $\hat{\beta}^{UE}$. For any estimator $\hat{\beta}^*$, the SRE of $\hat{\beta}^*$ with respect to $\hat{\beta}^{UE}$ is defined as

$$SRE\left(\hat{\beta}^{UE}, \hat{\beta}^*\right) = \frac{\text{SimulatedMSE}(\hat{\beta}^{UE})}{\text{SimulatedMSE}(\hat{\beta}^*)} = \frac{\text{Simulated } \sum_{i=1}^{p} \left(\beta_i - \beta_i^{UE}\right)^2}{\text{Simulated } \sum_{i=1}^{p} \left(\beta_i - \beta_i^{*}\right)^2}. \tag{19}$$

Keep in mind that an SRE is larger than the one that indicates the degree of superior of the estimator $\hat{\beta}^*$ over $\hat{\beta}^{UE}$.

3.1 Model with Correct Candidate Submodel ($\boldsymbol{\Delta^* = 0}$)

First, the case when the candidate submodel is assumed to be correct, is $\Delta^* = 0$. Various choices of active and inactive predictors are provided for $(p_1, p_2) = (4, 3), (4, 5), (4, 7), (4, 10)$, and $(4, 15)$, and the SRE results are reported in Tables 1, 2 and 3. The tuning parameter γ of the two penalty estimators is estimated using 10 fold-cross validation. The findings from Tables 1, 2 and 3 are summarized as follows:

We note that SREs of all the estimators increase as the number of inactive predictors $_2$ is increased for fixed λ and α. Interestingly, the restricted estimator is the best, and all estimators are superior to the unrestricted estimator for all configurations except the ridge regression estimator. The linear shrinkage estimator depends on the choice of λ. Its SRE decreases sharply to 1 as $\lambda \to 0$ and approaches to SRE of the restricted estimator for higher value of λ. The SREs of the pretest estimators depend on the size of the test α. For small α,

Table 1. The SREs of the estimators with respect to the UE for $\lambda = 0.25$ at $\Delta^* = 0$.

Estimator		Number of in active (p_2)				
		3	5	7	10	15
RE		1.693	2.104	2.735	3.544	6.120
LS		1.204	1.280	1.356	1.424	1.543
PT	$\alpha = 0.01$	1.631	2.002	2.542	3.162	4.558
	$\alpha = 0.05$	1.472	1.759	2.147	2.513	3.037
	$\alpha = 0.10$	1.369	1.606	1.804	2.108	2.398
	$\alpha = 0.35$	1.128	1.238	1.312	1.368	1.441
SP	$\alpha = 0.01$	1.190	1.263	1.335	1.396	1.489
	$\alpha = 0.05$	1.152	1.220	1.283	1.333	1.393
	$\alpha = 0.10$	1.125	1.187	1.227	1.278	1.325
	$\alpha = 0.35$	1.049	1.088	1.109	1.125	1.147
S		1.158	1.467	1.818	2.300	3.363
S^+		1.204	1.549	1.924	2.428	3.604
LASSO		1.226	1.293	1.470	1.819	2.383
Ridge		0.561	0.668	0.867	1.029	1.332

Table 2. The SREs of the estimators with respect to the UE for $\lambda = 0.50$ at $\Delta^* = 0$

Estimator		Number of in active (p_2)				
		3	5	7	10	15
RE		1.693	2.104	2.735	3.544	6.120
LS		1.416	1.610	1.839	2.069	2.556
PT	$\alpha = 0.01$	1.631	2.002	2.542	3.162	4.558
	$\alpha = 0.05$	1.472	1.759	2.147	2.513	3.037
	$\alpha = 0.10$	1.369	1.606	1.804	2.108	2.398
	$\alpha = 0.35$	1.128	1.238	1.312	1.368	1.441
SP	$\alpha = 0.01$	1.384	1.566	1.773	1.968	2.315
	$\alpha = 0.05$	1.298	1.454	1.622	1.765	1.954
	$\alpha = 0.10$	1.239	1.376	1.473	1.608	1.737
	$\alpha = 0.35$	1.089	1.162	1.206	1.240	1.286
S		1.158	1.467	1.818	2.300	3.363
S^+		1.204	1.549	1.924	2.428	3.604
LASSO		1.226	1.293	1.470	1.819	2.383
Ridge		0.561	0.668	0.867	1.029	1.332

Table 3. The SREs of the estimators with respect to the UE for $\lambda = 0.75$ at $\Delta^* = 0$

Estimator		Number of in active (p_2)				
		3	5	7	10	15
RE		1.693	2.104	2.735	3.544	6.120
LS		1.597	1.927	2.381	2.913	4.354
PT	$\alpha = 0.01$	1.631	2.002	2.542	3.162	4.558
	$\alpha = 0.05$	1.472	1.759	2.147	2.513	3.037
	$\alpha = 0.10$	1.369	1.606	1.804	2.108	2.398
	$\alpha = 0.35$	1.128	1.238	1.312	1.368	1.441
SP	$\alpha = 0.01$	1.546	1.848	2.245	2.669	3.556
	$\alpha = 0.05$	1.413	1.655	1.953	2.227	2.616
	$\alpha = 0.10$	1.326	1.53	1.689	1.925	2.159
	$\alpha = 0.35$	1.116	1.215	1.278	1.326	1.392
S		1.158	1.467	1.818	2.300	3.363
S^+		1.204	1.549	1.924	2.428	3.604
LASSO		1.226	1.293	1.470	1.819	2.383
Ridge		0.561	0.668	0.867	1.029	1.332

the pretest estimator is comparable to that of the restricted estimator. On the contrary, its SRE decreases as the size increases.

Moreover, the performance of the shrinkage pretest estimator depends on the choice of λ and α. Its performance becomes poorer with an increase in and a decrease in λ. Unsurprisingly, the pretest estimator outperforms the linear shrinkage estimator at $\Delta^* = 0$. As we would expect, the positive-part Stein-type shrinkage estimator outperforms the Stein-type shrinkage estimator in any situation.

Similar results are observed for the LASSO estimator which performs better than the estimator based on the Stein-type strategy for small p_2. On the other hand, the estimator based on the Stein-type strategy is preferable when there are many inactive predictors. However, the ridge regression estimator does not perform well, but its performance slowly improves as the number of inactive predictors increases.

3.2 Model with Correct and Incorrect Candidate Submodel ($\boldsymbol{\Delta^* \geq 0}$)

The penalty estimators are not included in the $\Delta^* > 0$ case because these estimators do not take advantage of the fact that the regression parameter lies in the subspace $\beta_2 = 0$. In this case, the simulation model has an active coefficient vector $\beta_1 = (0.5, -1.5, 0.2)$ and inactive coefficient vector $\beta_2 = (\beta_4, a)$ where β_4 is a scalar and assumes various values; thus, $\Delta^* = (\beta_4)^2$ when the candidate submodel is used. We choose $\beta_4 = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.8, 1.0, 1.5,$ and 2.0

and a is $k \times 1$ zero vector with different dimensions, $k = p_2 - 1$ to be the number of inactive predictors in the model and we use $p_2 = 5, 7, 10$, and 15 in the simulations.

The choice of α was fixed to be 0.01, 0.05, 0.10, $\lambda = 0.75$ and $n = 250$. SREs of the proposed estimators are are reported in Table 4 are graphically represented in Figs. 1, 2 and 3. The findings are summarized as follows:

The restricted estimator $\hat{\beta}^{RE}$ outperforms all the estimators at and near $\Delta^* = 0$. In contrast, as Δ^* becomes larger than zero, the relative efficiency of $\hat{\beta}^{RE}$ as well as the linear shrinkage estimator decrease and become unbounded. However, the relative efficiency of all the other estimators remains bounded and approaches to 1. These show that an incorrect candidate submodel is fatal to the restricted and linear shrinkage estimators.

The pretest estimator is much better than the shrinkage pretest estimator when the candidate submodel is correct, i.e. $\Delta^* = 0$. On the contrary, the shrinkage pretest estimator does well relative to the pretest estimator in a small part of the parameter space. However, the SRE of both pretest and shrinkage pretest estimators approaches to one as Δ^* moves away from zero, but after becoming inferior to $\hat{\beta}^{UE}$, and later at some point, they join the SRE of one from below. In addition, they outshine the estimator based on Stein-type strategy where Δ^* is near zero and for small and moderate p_2. The estimators based on Stein-type strategy are superior to the unrestricted estimator in the entire range of Δ^* especially their gain in risk reduction is impressive as p_2 increases. Lastly, we found that the estimators based on Stein-type strategy perform better than all other estimators in the wider range of Δ^* and these estimators are little impacted by severe departure from the restriction.

4 Real Data Example: South African Heart Disease Data

In this section, we apply the proposed estimators to the South African heart disease data set. Rousseauw et al. [14] described a retrospective sample of males in a heart-disease high-risk region of the Western Cape, South Africa. This study comprised over 462 samples and the set of variables is described in Table 5.

We notice that the condition index (CI) value is calculated as 392.718 which implies the existence of multicollinearity in this data set. After applying the variable selection procedure based on AIC criterion, BIC criterion, and LASSO, the results are given in Table 6.

Table 6 shows that the candidate submodel based on AIC and BIC criteria contains 5 active predictors, while LASSO selection procedure contains 7 active predictors. Hence, we will consider the candidate submodel with 5 active predictors that is tobacco, famhist, ldl, typea, and age. The restricted subspace is $\beta_2' = (\beta_{\text{adiposity}}, \beta_{\text{obesity}}, \beta_{\text{alcohol}}, \beta_{\text{sbp}}) = (0, 0, 0, 0), p = 9, p_1 = 5$, and $p_2 = 4$.

To examine the performance of the proposed estimators for the candidate submodel, we draw $m = 250$ bootstrap rows with replacement $N = 1,000$ times from the data. The performance of the proposed estimators with respect to the

Table 4. The SREs of RE, LS, PT, SP, S, and $S+$ with respect to the UE for $\lambda = 0.75$

p_2	Δ^*	RE	LS	PT			SP			S	S+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
5	0.0	2.581	2.311	2.428	2.057	1.806	2.195	1.909	1.705	1.597	1.726
	0.1	1.328	1.611	1.141	1.046	1.016	1.299	1.139	1.082	1.279	1.309
	0.2	0.884	1.227	0.836	0.882	0.907	0.976	0.957	0.960	1.187	1.195
	0.3	0.666	0.988	0.773	0.873	0.920	0.882	0.928	0.954	1.150	1.152
	0.4	0.54	0.833	0.795	0.910	0.958	0.879	0.947	0.975	1.126	1.127
	0.5	0.452	0.716	0.863	0.948	0.968	0.917	0.967	0.980	1.111	1.111
	0.8	0.309	0.515	0.958	0.988	0.996	0.973	0.992	0.998	1.088	1.088
	1.0	0.255	0.434	0.984	1.000	1.000	0.989	1.000	1.000	1.081	1.081
	1.5	0.184	0.322	1.000	1.000	1.000	1.000	1.000	1.000	1.074	1.074
	2.0	0.152	0.269	1.000	1.000	1.000	1.000	1.000	1.000	1.070	1.070
7	0.0	3.065	2.66	2.873	2.477	2.079	2.523	2.232	1.924	1.959	2.147
	0.1	1.701	1.962	1.406	1.215	1.147	1.550	1.286	1.197	1.499	1.540
	0.2	1.183	1.548	1.000	0.973	0.975	1.140	1.042	1.021	1.349	1.359
	0.3	0.909	1.29	0.876	0.927	0.955	0.985	0.978	0.984	1.282	1.284
	0.4	0.748	1.116	0.857	0.939	0.963	0.938	0.971	0.981	1.240	1.241
	0.5	0.643	0.993	0.891	0.962	0.979	0.944	0.979	0.989	1.217	1.217
	0.8	0.449	0.735	0.969	0.991	0.996	0.981	0.994	0.997	1.169	1.169
	1.0	0.376	0.632	0.989	0.997	1.000	0.993	0.998	1.000	1.155	1.155
	1.5	0.278	0.483	1.000	1.000	1.000	1.000	1.000	1.000	1.139	1.139
	2.0	0.222	0.394	1.000	1.000	1.000	1.000	1.000	1.000	1.129	1.129
10	0.0	4.693	3.702	3.987	3.087	2.396	3.277	2.680	2.175	2.612	2.942
	0.1	2.43	2.661	1.982	1.531	1.366	2.124	1.604	1.414	1.944	2.048
	0.2	1.669	2.124	1.301	1.135	1.074	1.499	1.230	1.135	1.686	1.730
	0.3	1.263	1.755	1.029	0.991	0.983	1.197	1.079	1.037	1.569	1.584
	0.4	1.024	1.510	0.911	0.931	0.949	1.048	0.997	0.99	1.485	1.49
	0.5	0.868	1.335	0.857	0.926	0.955	0.974	0.974	0.983	1.433	1.434
	0.8	0.610	1.001	0.890	0.959	0.980	0.942	0.977	0.989	1.326	1.327
	1.0	0.508	0.858	0.928	0.980	0.992	0.959	0.988	0.995	1.300	1.300
	1.5	0.368	0.648	0.996	1.000	1.000	0.998	1.000	1.000	1.262	1.262
	2.0	0.295	0.529	1.000	1.000	1.000	1.000	1.000	1.000	1.236	1.236
15	0.0	6.502	4.665	4.986	3.510	2.651	3.879	2.976	2.372	3.58	3.961
	0.1	3.570	3.566	2.568	1.735	1.462	2.570	1.742	1.469	2.569	2.662
	0.2	2.466	2.926	1.575	1.258	1.164	1.703	1.310	1.197	2.163	2.196
	0.3	1.906	2.506	1.219	1.083	1.046	1.330	1.132	1.075	1.949	1.956
	0.4	1.549	2.188	1.062	1.018	1.011	1.152	1.053	1.029	1.818	1.819
	0.5	1.323	1.951	0.993	0.993	0.992	1.057	1.017	1.005	1.712	1.713
	0.8	0.927	1.503	0.971	0.992	0.997	0.991	0.997	0.999	1.560	1.560
	1.0	0.793	1.331	0.984	0.995	0.998	0.993	0.998	0.999	1.500	1.500
	1.5	0.574	1.015	0.999	0.999	1.000	0.999	0.999	1.000	1.423	1.423
	2.0	0.458	0.827	1.000	1.000	1.000	1.000	1.000	1.000	1.380	1.380

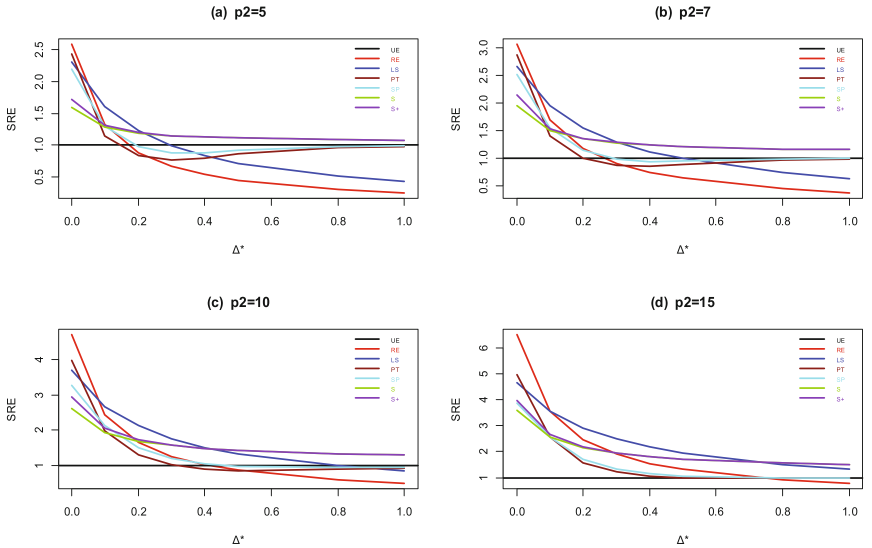


Fig. 1. SREs of RE, LS, PT, SP, S, and S^{+} with respect to the UE when the candidate subspace misspecifies β_4 as zero as of $\Delta^{*} = (\beta_4)^2$. Here, $p_1 = 3$, $\alpha = 0.01$.

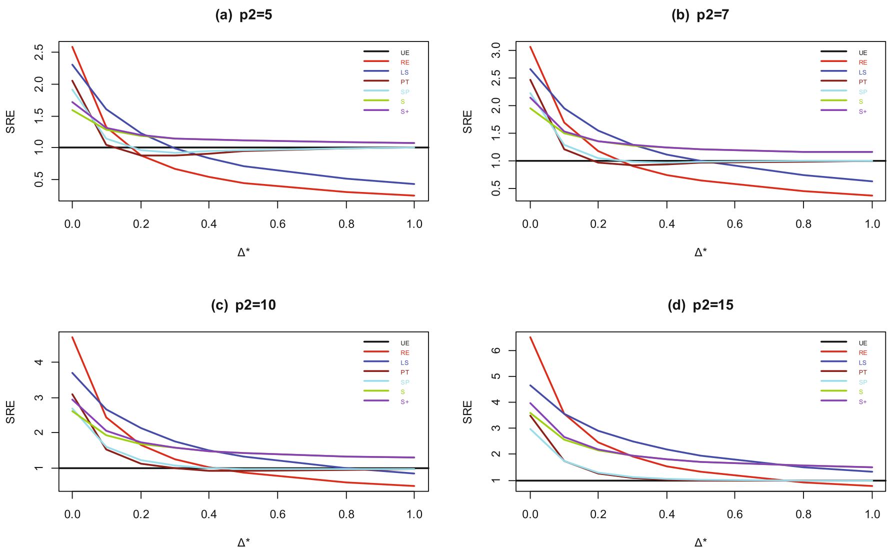


Fig. 2. SREs of RE, LS, PT, SP, S, and S+ with respect to the UE when the candidate subspace misspecifies β_4 as zero as of $\Delta^{*} = (\beta_4)^2$. Here, $p_1 = 3$, $\alpha = 0.05$.

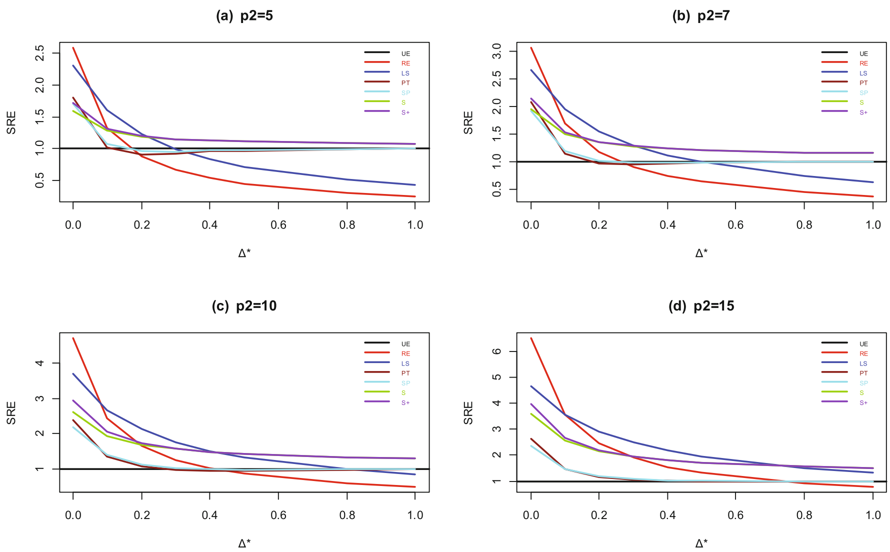


Fig. 3. SREs of RE, LS, PT, SP, S, and S+ with respect to the UE when the candidate subspace misspecifies β_4 as zero as of $\Delta^* = (\beta_4)^2$. Here, $p_1 = 3$, $\alpha = 0.10$.

Table 5. List of variables

Variable	Description
Response variable	
Chd	Coronary heart disease
Predictor variable	
Tobacco	Cumulative tobacco (kg)
Famhist	Family history of heart disease, a factor with levels absent and present
Ldl	Low density lipoprotein cholesterol
Typea	Type-A behavior
Age	Age at onset
Adiposity	Adiposity
Obesity	Obesity
Alcohol	Current alcohol consumption
Sbp	Systolic blood pressure

Table 6. Full and candidate sub-models for South African heart disease data

Selection criterion	Response	Active predictor
Full model	Chd	tobacco, famhist, ldl, typea, age, adiposity, obesity, alcohol, sbp
LASSO	Chd	tobacco, famhist, ldl, typea, age, obesity, sbp
AIC,BIC	Chd	tobacco, famhist, ldl, typea, age

unrestricted estimator is evaluated by simulated relative efficiency (SRE) and simulated relative prediction error (SPE) of the estimators which is defined as:

$$\text{SRE}\left(\hat{\beta}^{UE}, \hat{\beta}^{*}\right) = \frac{\text{Simulated } \sum_{i=i}^{p}\left(\beta_i^{\text{true}} - \hat{\beta}_i^{UE}\right)}{\text{Simulated } \sum_{i=1}^{p}\left(\beta_i^{\text{true}} - \hat{\beta}_i^{*}\right)}$$

$$\text{and SPE}\left(\hat{\beta}^{UE}, \hat{\beta}^{*}\right) = \frac{\text{Simulated } \sum_{i=1}^{m}\left(Y_i - \pi(\mathbf{x}'_i\hat{\beta}^{UE})\right)}{\text{Simulated } \sum_{i=1}^{m}\left(Y_i - \pi(\mathbf{x}'_i\hat{\beta}^{*})\right)}, i = 1, 2, \cdots, m.$$

Note that SPE is less than one; this means the unrestricted estimator is doing better. This study assumed the empirical distribution $\hat{F}$ based on 462 actual observations to be the true distribution and the resulting logistic regression coefficient $\hat{\beta}$'s to be the true parameter values. We assumed $\alpha = 0.01$ and $\lambda = 0.50$. The results of the point estimates, standard errors, SREs, and SPEs of the estimators are shown in Table 7.

Table 7 reveals that the restricted estimator is the best, and all the estimators outperform the unrestricted estimator. The performance of linear shrink-

Table 7. Estimates (first row) and standard errors (second row) of the coefficients for active predictors. The SRE and SPE columns give the relative efficiency and relative prediction error of the estimators with respect to UE, respectively.

Estimator	β_{tobacco}	β_{famhist}	β_{ldl}	β_{typea}	β_{age}	SRE	SPE
UE	0.084	0.948	0.186	0.042	0.048	1.000	1.000
	0.041	0.340	0.087	0.018	0.018		
RE	0.083	0.917	0.167	0.038	0.052	2.156	1.364
	0.039	0.324	0.077	0.017	0.015		
LS	0.085	0.933	0.176	0.040	0.050	1.653	1.360
	0.040	0.330	0.079	0.018	0.015		
PT	0.084	0.922	0.170	0.039	0.051	1.836	1.230
	0.039	0.329	0.079	0.018	0.016		
SP	0.085	0.935	0.178	0.040	0.050	1.507	1.274
	0.040	0.333	0.081	0.018	0.016		
S	0.084	0.938	0.180	0.041	0.049	1.309	1.156
	0.040	0.334	0.082	0.018	0.016		
S^{+}	0.084	0.938	0.180	0.041	0.049	1.361	1.191
	0.040	0.334	0.082	0.018	0.016		
LASSO	0.071	0.788	0.147	0.029	0.042	1.360	1.051
	0.038	0.353	0.083	0.020	0.014		
Ridge	0.072	0.761	0.146	0.028	0.035	1.552	1.210
	0.030	0.311	0.068	0.018	0.015		

age, pretest, and shrinkage pretest estimators is dominated by the restricted estimator as our data in the resampling scheme are generated from an empirical distribution where the candidate submodel is correct, that is $\Delta^* = 0$.

Furthermore, the positive-part Stein-type shrinkage estimator outshines the LASSO estimator when there are moderate or relatively large numbers of the inactive predictor in the model. In fact, the true parameter values are not exactly zero, and there is the multicollinearity problem. Unsurprisingly, the ridge regression estimator performs well. Lastly, the LASSO estimator shows good performance in terms of SRE but not in terms of SPE because of the instability estimation.

5 Discussion and Conclusions

In this article, we compared various estimators based on pretest and Stein-type strategy and two penalty estimators to the unrestricted and restricted maximum likelihood estimators in the context of the logistic regression model under the restriction of parameter. We established the properties of the proposed estimators via Monte Carlo simulation study.

By using Monte Carlo simulation study, we found that the performance of a restricted maximum likelihood estimator depends heavily on the quality of the candidate submodel. The restricted maximum likelihood estimator is the best estimator when the candidate submodel is correct or nearly correct. For any scenario, the estimators based on Stein-type strategy outperform the unrestricted maximum likelihood in the entire parameter space, especially the truncated version of the Stein-type shrinkage estimator. On the contrary, the performance of the estimators based on the preliminary test procedure lacks this property. The LASSO estimator is preferable to the Stein-type estimators when the number of inactive predictors is small. In contrast, the Stein-type estimators dominate the LASSO estimator only when the number of inactive predictors is relatively large. The ridge regression cannot produce a parsimonious model, thus the ridge regression estimator does not perform well when the model is sparse. However, the positive-part Stein-type shrinkage estimator is robust. It would be interesting, therefore, to investigate the relative performances of adaptive LASSO and smoothly clipped absolute deviation (SCAD) estimators. We will leave this for further consideration.

Acknowledgements. We would like to thank all the referees, the Editor, and the Associate Editor for their valuable suggestions on the revision of the article.

References

1. Ahmed SE (1992) Shrinkage preliminary test estimation in multivariate normal distributionsm. J Stat Comput Simul 43(3–4):177–195
2. Ahmed SE, Amezziane M (2016) Shrinkage-based semiparametric density estimation. Statistical Methodology

3. Ahmed SE, Raheem E, Hossain S (2011) Absolute Penalty Estimation. Springer International Encyclopedia of statistical science, New York
4. Ahmed SE, Yüzbaşı B (2016) Big data analytics: integrating penalty strategies. International Journal of Management Science & Engineering Management 11(2016):105–115
5. Al-Momani M, Hussein AA, Ahmed SE (2016) Penalty and related estimation strategies in the spatial error model: estimation strategies in sem model
6. Gao X, Ahmed SE, Feng Y (2016) Post selection shrinkage estimation for high-dimensional data analysis. Appl Stoch Models Bus Ind
7. Gourieroux CS, Monfort A (1981) Asymptotic properties of the maximum likelihood estimator in dichotomous logit models. J Econometrics 17(1):83–97
8. Hilbe JM (2009) Logistic Regression Models. Chapman and Hall, Boca Raton
9. Hoerl AE, Kennard RW (2000) Ridge regression: biased estimation for nonorthogonal problems. Technometrics 42(1):80–86
10. Hosmer DW, Lemeshow S (2000) Applied Logistic Regression. Wiley
11. Hossain S, Ahmed S (2014) Shrinkage estimation and selection for a logistic regression model
12. Hossain S, Ahmed SE, Doksum KA (2014) Shrinkage, pretest, and penalty estimators in generalized linear models. Stat Methodol 24(2):52–68
13. Lisawadi S, Shah MKA, Ahmed SE (2016) Model selection and post estimation based on a pretest for logistic regression models. J Stat Comput Simul 86(17):1–17
14. Rossouw JE, Du PJ, et al (1983) Coronary risk factor screening in three rural communities. The CORISs baseline study. S Afr Med J 64(12):430–436
15. Tibshirani R (2011) Regression shrinkage and selection via the lasso. J Roy Stat Soc 58(3):267–288
16. Yüzbaşı B, Ahmed SE (2015) Shrinkage ridge regression estimators in high-dimensional linear models. Springer, Heidelberg
17. Yüzbaşı B, Ahmed SE (2016) Shrinkage and penalized estimation in semi-parametric models with multicollinear data. J Stat Comput Simul

Multi-objective Job Shop Rescheduling with Estimation of Distribution Algorithm

Xinchang Hao[1(✉)], Lu Sun[2], and Mitsuo Gen[3]

[1] Waseda University, Tokyo 163-8001, Japan
haoxc@live.com
[2] Dalian University of Technology, Dalian 116024, People's Republic of China
[3] Fuzzy Logic Systems Institute, Tokyo University of Science, Tokyo 163-8001, Japan

Abstract. To solve the moJSRP model, with the framework of proposed MoEDA, the probability model of the operation sequence is estimated firstly. For sampling the processing time of each operation with the Monte Carlo methods, allocation method is used to decide the operation sequence, and then the expected makespan and total tardiness of each sampling are evaluated. Subsequently, updating mechanism of the probability models is proposed according to the best solutions to obtain. Finally, for comparing with some existing algorithms by numerical experiments on the benchmark problems, we demonstrate the proposed effective estimation of distribution algorithm can obtain an acceptable solution in the aspects of schedule quality and computational efficiency.

Keywords: Job shop rescheduling · Multi-objective optimization · Estimation of distribution algorithms

1 Introduction

In current manufacturing systems, production processes and management are involved in many unexpected events and new requirements are emerging constantly. For modern flexible decision making in manufacturing systems it is necessary to adapt the unexpected changes with rescheduling processes. As a result, production managers in the volatile environment have to generate high quality schedules [3,8]. Rescheduling is defined as periodically or event-driven operation to better arranging activities for the next time period based on both the state of the system and the vested schedule. Many literatures have reported researches on the rescheduling topic. Typically, Vieira et al. presented a comprehensive survey for most applications of rescheduling in manufacturing systems [8]. This methodology is summarized in Table 1.

In this paper, job shop rescheduling problem (JSRP) in dynamic environment will be studied, more specifically, the study takes into account the effect of random job arrivals and machine breakdowns. Since JSRP is ubiquitous in manufacturing systems, it has attracted great attention to design optimization methods

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_3

Table 1. Rescheduling framework

Rescheduling environments	Static (finite set of jobs)	Deterministic (all information given)	
		Stochastic (some information uncertain)	
	Dynamic (infinite set of jobs)	No arrival variability (cyclic production)	
		Arrival variability (flow shop)	
		Process flow variability (job shop)	
Rescheduling strategies	Dynamic (no schedule)	Dispatching rules	
		Control-theoretic	
	Predictive-reactive (generate and update)	Rescheduling policies	Periodic
			Event-driven
			Hybrid
Rescheduling methods	Schedule generation	Nominal schedules	
		Robust schedules	
	Schedule repair	Right-shift rescheduling	
		Partial rescheduling	
		Complete regeneration	

for finding the desired effective solutions. It is well-known that Job shop scheduling problem is NP-hard, and various methods have been introduced to obtain optimum solutions. From the practical point of view, many dispatching rules and composite dispatching rule have been developed for the JSRP [8]. Sabuncuoglu and Kizilisik proposed several reactive scheduling policies and tested their performances with various experimental conditions, processing time variations, and random machine breakdown Table 1. Sha and Liu presented extended data mining tools for extracting knowledge of job scheduling with respect to due date assignment in a dynamic job shop environment [7]. Vinod and Sridharan presented the salient aspects of a simulation-based experimental study on scheduling rules in a dynamic job shop where the setup times are sequence dependent [9]. However, a key challenge is to balance available capacities among jobs at different processing stages, especially in the shops with re-entrant flow, like those in semiconductor wafer fabrication plants [5]. Furthermore, in job shop rescheduling, job arrival of specific product is often unexpected, critical job information is not available in advance which would eliminate schedulers to anticipate future workloads effectively. However, all these researches transform the multi-objective problem to a single objective problem using a weighting method. They do not discuss the implications of non-dominated solutions for the multi-objective dynamic job shop scheduling problems.

In this paper, to solve the moJSRP model, with the framework of proposed MoEDA, the probability model of the operation sequence is estimated firstly. For sampling the processing time of each operation with the Monte Carlo methods, allocation method is used to decide the operation sequence, and then the expected makespan and total tardiness of each sampling are evaluated. Subsequently, updating mechanism of the probability models is proposed according to the best solutions to obtain.

The rest of the paper is organized as follows: in Sect. 2, the mathematical model of moJSRP is formulated; Sect. 3 describes the proposed EDA-based approach for solving the moJSRP; in Sect. 4, the computational experiments and analysis of results are presented; finally, Sect. 5 gives the summary and the conclusions of this research.

2 Mathematical Formulation

(1) Assumption

The job shop scheduling problem (JSP) concerns with the determination of the operation sequences on the machines so that the makespan is minimized. It may consist of several assumptions as follows:

A1. Each machine processes only one job at a time.
A2. Each job is processed on one machine at a time.
A3. Jobs never visit the same machine twice.
A4. For each operation, the processing time and machine assignment are known in advance.
A5. There are no precedence constraints among operations of different jobs.
A6. Operations cannot be interrupted.
A7. Neither release times nor due dates are specified.

For a new job arrival, the following assumptions are given:

A8. The 1rst operation of the new job has to be started after the processing operation is finished on one machine.
A9. For new job, the processing time and machine assignment of each operation are determinate.
A10. There are no precedence constraints between operations of original jobs and new job.
A11. For new job, neither release times nor due dates are specified.

The JSP has already been confirmed as one of the NP-hard combinatorial problems. There are N jobs and M machines to be scheduled; furthermore, each job is composed of a set of operations and the operation order on machines is pre-specified. Each operation is characterized by the required machine and the fixed processing time [3].

(2) Notation

Indices:

i, i', k: job index; where i is original job index, i' is any job index, k is the index of new arrival job;
j, j', l: operation index; where j is operation index of original job, j' is any operation index of any job, and l is the operation index of new arrival job.

Parameters:

J: number of original jobs;
N: number of new jobs;
M: number of machines;
N_i: number of operations of original job i;
N_k: number of operations of new job k;
o_{ij}: operation j of original job i;
o_{kl}: operation l of new job k;
$o_{i'j'}$: operation j' of any job i';
m_{ij}: machine on which operation is to be assigned;
p_{ij}: processing time of operation o_{ij};
p_{kl}: processing time of operation o_{kl};
t_ij: begin time of original operation o_{ij};
t_b: predefined time at which schedule is renewed, i.e., rescheduling point'.

Decision Variables:

t'_{ij}: begin time of operation o_{ij} in the original schedule;
t'_{kl}: begin time of operation o_{kl} in the new schedule.

The first objective Eq. (1) of the moJSRP is to minimize the makespan of new job. The second objective Eq. (2) is to minimize the costs caused by disruption on original schedule, where the term Δt_{ij} and $\Delta\bar{t}$ denote the deviation and its mean of begin time respectively in the original schedule.

$$\min \ t_{MS} = \max_k (t'_{kN_k} p_{kN_k}) \tag{1}$$

$$\min \ \sigma = \frac{1}{\sum_i N_i} \sum\nolimits_i \sum\nolimits_j (\Delta t_{ij} - \Delta\bar{t})^2 \tag{2}$$

$$\text{where } \Delta t_{ij} = \left| t_{ij} - t'_{ij} \right| \tag{3}$$

$$\Delta\bar{t} = \frac{1}{\sum_i N_i} \sum\nolimits_i \sum\nolimits_j |\Delta t_{ij}| \tag{4}$$

$$\text{s. t.} \ \left(t'_{t,l-1} + p_{k,i-1}\right) \leq t'_{kl}, \forall k, l > 1 \tag{5}$$

$$(t_{i'j'} - (t_{ij} + p_{ij})) \wedge (t_{ij} - (t_{i'j'} + p_{i'j'})) + \psi \left| m_{ij} - m_{i'j'} \right| \geq 0, \forall o_{ij}, o_{i'j'} \tag{6}$$

$$t'_{ij} = t_{ij}, t_{ij} \leq t_b \tag{7}$$

$$t'_{ij} > 0, \forall i, j \tag{8}$$

$$t'_{kl} > 0, \forall k, l, \tag{9}$$

where, inequity (5) ensures that the $(l-1)^{th}$ operation of job k should be processed before the $l-th$ operation in the same job. Equation (6) expresses the operation precedence constraints. Equation (7) ensures that the original operations that are fixed for execution at the rescheduling point cannot be changed. Inequities (8) and (9) represent the nonnegative restrictions.

3 Effective Multiobjectives Estimation of Distribution Algorithm

3.1 A Descriptive Analysis of Air Pollution and Death Data

The multiple objective optimization problems have been receiving growing interest from researchers with various backgrounds since early 1960. There are a number of scholars who have made significant contributions to the problem. In the paper, we present a scheme of hybrid multiobjective estimation of distribution algorithm (h-MoEDA). In h-MoEDA, MoEDA takes care of exploration that tries to identify the most promising search space regions, and modeling the distribution of the stochastic variables. For NP-hard combinatorial optimization problems, it became necessary that the global search dynamics of EDAs ought to be complemented with local search refinement [6]. Variable neighborhood search (VNS) is proposed to improve the solution quality, and the simple principle for driving this improvement is based on the systematic change of neighborhood within a possibly randomized local search.

In Fig. 1, h-MoEDA starts by generating the solutions kept in the population randomly. The set of solutions of high fitness, which refers to promising solutions, is selected from the population using a selection method, and the promising solutions are used to learn the probability model. Thereafter, probabilities defined by the probability model are estimated, and the new candidate solutions

Hybrid moEDA routine

Input:
X: $X = (X_1, X_2, \ldots, X_n)$, vector of the decision variables.
$popSize$: number of the population solutions kept by moEDA.
gen : number of the generations (iterations).
$prRate$: rate of the promising solutions kept by moEDA.
$elimRate$: elimination rate of the keeping promising solutions.
Output:
S_{best}: the best Pareto solution.

begin
Initialization:
Step 1: $t \leftarrow 0$.
Step 2: Initialize the population $Pop(0)$ with the size $popSize$ solutions randomly.
Step 3:Calculate objectives $f_k(Pop), k = 1, \ldots, q$ by decoding routine.
Step 4 Create Pareto optimal solutions $E(Pop)$ by nondominated routine.
Step 5 Calculate fitness $eval(Pop)$ by HSS-EA routine and keep best Pareto solution S_{best}.
Optimization:
while *terminating criteria are not met* **do**
Step 6: Select the promising data set D from $Pop(t)$ consisting of $popSize \times prRate$ solutions.
Step 7: Estimate the probabilities $p(x_i)$ for each variable X_i over the set D, and sample $candidates$ based on the $p(x_i)$.
Step 8: Perform a problem-specific local search $DoLocalSearch(candinates)$.
Step 9:Calculate objectives $f_k(Pop), k = 1, \ldots, q$ by decoding routine.
Step 10:update Pareto optimal solutions $E(Pop)$ by nondominated routine.
Step 11 Calculate fitness $eval(Pop)$ by HSS-EA routine and keep best Pareto solution S_{best}.
Step 12: Replace $popSize \times elimRate$ worse items of the keeping population solutions with the best solutions $candidates$.
Step 13: $t \leftarrow t + 1$.
end
end

Fig. 1. Pseudo-code for h-MoEDA

are sampled according to the given sampling method. For a candidate solution, problem specific local search algorithm is used to improve each candidate solution to reach a local optimum. Finally, the new solutions are evaluated, and these with high fitness are incorporated into a solution pool, which keeps these individuals contributing to the makeup of promising solutions. The iteration will continue until the predefined termination criteria are met.

The following subparagraph presents the vital components of the proposed h-MoEDA in detail. Firstly, the representation of a chromosome for an individual is described. Then, transition probability model is presented. Finally, fitness assignment mechanisms are discussed.

3.2 Solution Representation

The representation of the operation sequence uses job-based encoding [1], and the length of the chromosome equals the total number of operations. The job number denotes the operation of each job, and the l^{th} occurrence of a job number refers to the l^{th} operation in the sequence of this job. Figure 1 presents an example considering a rescheduling with a 4 jobs and 3 machines.

ID	1	2	3	4	5	6	7	8	9	10	11	12
Operation	o_{11}	o_{12}	o_{13}	o_{21}	o_{22}	o_{23}	o_{31}	o_{32}	O_{33}	o_{41}	o_{42}	O_{43}
Chromosome v	-	4	7	-	-	3	-	-	6	1	2	5

Fig. 2. Illustration of the representation of a solution for moJSRP

Giffler-Thompson based heuristic algorithm [4] based decoding procedure is applied to generate a feasible schedule. This produces a timetable with the start and end of the processing period, and the makespan time. The Gantt chart of this solution is illustrated in Fig. 3.

ID	1	2	3	4	5	6	7	8	9	10	11	12
Operation	o_{11}	o_{12}	o_{13}	o_{21}	o_{22}	o_{23}	o_{31}	o_{32}	O_{33}	o_{41}	o_{42}	O_{43}
Chromosome v	-	4.37	7.69	-	-	3.91	-	-	6.35	1.01	2.67	5.93

Fig. 3. Gantt chart of the schedule encoded by the operation sequence shown in Fig. 2

An example of reschedule repaired priority-based chromosome are shown in Fig. 4.

3.3 Transition Probability

For traditional EAs, the representation of a chromosome for an individual is generated by mapping the decision space into the search space or directly encoding the decision variables. Each position in a probability vector indicates the

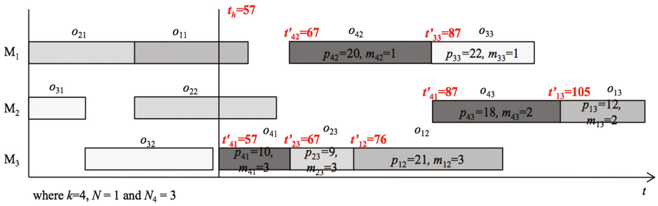


Fig. 4. An illustration of reschedule generated by repaired priority-based representation

distribution of probability regarding each variable. When prior knowledge of distribution is not assumed, the domain of discrete variable X is a set of pre-defined values (x). The distribution of random variable X has the same equal probability; the initialization is as following:

$$p_{t=0}(X) = 1/|X|, \tag{10}$$

where $|X|$ denotes the number of values in the set of domain X.

After the initialization of EDA, probability sample for new alternative solutions and the new solutions are evaluated according to a specific system objective. EDA collects all new alternative solutions and replaces the inferior solutions in the promising data. The probability distribution of X can be estimated as follows:

$$B\_t(X = x) = (N(X = x) + 1/|X|)/\text{prSize} + 1/|X|, \tag{11}$$

where $N(X = x)$ denotes the number of instances in promising solutions with variable $X = x$, and represents the low bound to the probability of X.

The distribution probability of X in the probability vector is learned toward the estimated distribution of promising data, as follows:

$$P_{t+1}(X = x) = (1 - \alpha) P_t(X = x) + \alpha B_t(X = x), \tag{12}$$

where α denotes the learning rate from the current promising solutions; in particular, for $\alpha = 1$, the probability distribution is completely reconstructed by the current promising solutions.

To maintain the diversity of sampling, the distribution probability of X is updated toward the estimation distribution. The distribution can be tuned with probability pm of the mutation, and the mutation is performed using the following definition:

$$P_{t+1}(X = x) = \frac{P_t(X = x) + \lambda_m}{\sum_{x' \in X \setminus \{x\}} \max\left(P_t(x) - \frac{\lambda_m}{|X|-1}, \varepsilon\right) + (P_t(X = x) + \lambda_m)}, \tag{13}$$

where λ_m is the mutation shift that controls the amount for mutation operation, and ε is a small probability value to avoid the negative probability value.

3.4 Fitness Assignment Mechanisms

The multiple objective optimization problems have been receiving growing interest from researchers with various backgrounds since early 1960. There are a number of scholars who have made significant contributions to the problem. Among them, Pareto is perhaps one of the most recognized pioneers in this field and recently. Many multi-objective genetic algorithms differ mainly in the fitness assignment strategy which is known as an important issue in solving multiple objectives optimization problem [1]. Zhang et al. proposed a hybrid sampling strategy-based evolutionary algorithm (HSS-EA) in which a Pareto dominating and dominated relationship-based fitness function (PDDR-FF) is used to evaluate the individuals. PDDR-FF of an individual S_i is calculated by the following fitness assignment function (14):

$$\text{eval}\,(S_i) = q\,(S_i) + 1/\,(p\,(S_i) + 1)\; i = 1, 2, \cdots, \text{popSize}, \tag{14}$$

where $q(\cdot)$is the number of individuals which can dominate the individual $S \cdot p(\cdot)$ is the number of individuals which can be dominated by the individual S_i. The PDDR-FF can set the obvious difference values between the nondominated and dominated individuals as shown in Fig. 5.

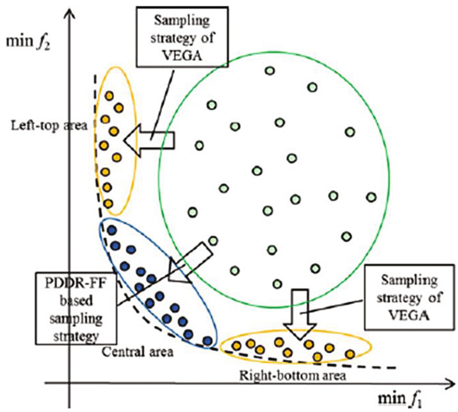


Fig. 5. The description of HSS-EA

4 Experiments and Discussion

To examine the practical viability and efficiency of the pro- posed MoEDA, we designed a numerical study to com- pare MoEDA with efficient algorithms from previous studies. The proposed MoEDA was compared with adaptive weight genetic algorithm (awGA) [1], Non-dominated Sorting Genetic Algorithm II (NSGA-II) [2] and Strength Pareto Evolutionary Algorithm 2 (SPEA2) [10] for a set of simulation data of testing standard benchmark problems. All of the above algorithms were implemented using JAVA under the Eclipse environment, and the simulation experiments were conducted on Intel Core $i5$ (2.3 GHz clock) with 4G memory. Data were collated from 30 test runs for each algorithm. In order to compare the performance of these algorithms fairly and under the same environment, the strategies of related algorithms and their respective parameters are presented in Table 2.

Table 2. The parameters of NSGA-II, SPEA2, AWGA and h-MoEDA

	NSGA-II, SPEA2 and awGA	MoEDA
Interaction	1000	1000
Population	200	200
Selection	Roulette	Tournament (k)
Operators	Crossover (P_c)	Sampling
	Mutation (P_m)	Improvement (P_i)
Parameters	$P_m = 0.20$	PromisingRate $= 0.7$
	$P_c = 0.80, w_2 = 0.6$	$\alpha = 0.02, \beta = 0.02$
	$w_1 = 0.4, w = 0.6$	$P_m = 0.4, k = 2$

In order to evaluate the performance of a given algorithm for MoJSRP, the following two performance measures are considered in this paper. Let S_j be a solution set for each solution method ($j = 1, 2, 3, 4$). $S*$ is a known set of the Pareto-optimal set.

Coverage $C(S_1, S_2)$ is the percent of the individuals in S_2 who are weakly dominated by S_1. The value $C(S_1, S_2) = 1$ means that all individuals in S_2 are weakly dominated by S_1.

On the contrary, $C(S_1, S_2) = 0$ denotes that none of the individuals in S_2 is weakly dominated by S_1 [10]. The larger $C(S_1, S_2)$ is; the better S_1 outperforms S_2 in C.

4.1 Convergence

Table 4 shows a comparison of coverage by MoEDA, awGA, NSGA-II and SPEA2. Figure 6 presents the distribution range of coverage conducted on LA35. From Table 3 and Fig. 6, it is easy to see that the h-MoEDA is better than MSGA-II, SPEA2 and awWA on C measure. Such better convergence should mainly attribute to the hybrid sampling strategy of VEGA's preference for the edge region of the Pareto front and PDDR-FF's tendency converging toward the centre area of the Pareto front.

They preserve better performances both in efficacy and efficiency. Especially, h-MoEDA can also keep diversity evenly without special distribution mechanisms like NSGA-II and SPEA2. h-MoEDA is better than MSGA-II, SPEA2

Table 3. Comparison of coverage measure by h-MoEDA, AWGA, NSGA-II, SPEA2 and AWGA

Problem	Algorithm	Mean		Improved	SD	
		C(h-MoEDA,A)	C(A, h-MoEDA)	(C(h-MoEDA,A)-C(A,h-MoEDA))(%)	C(h-MoEDA,A)	C(A, h-MoEDA)
LA29	NSGA-II	0.4669	0.3524	11.45	0.0842	0.0709
	SPEA2	0.4965	0.3224	17.41	0.0696	0.0603
	awGA	0.5731	0.2541	31.90	0.0451	0.0553
LA35	NSGA-II	0.6846	0.2824	36.63	0.0777	0.0794
	SPEA2	0.6606	0.2618	39.89	0.0860	0.0631
	awGA	0.6844	0.2127	46.73	0.0672	0.0460
LA38	NSGA-II	0.4543	0.3425	11.19	0.0779	0.0754
	SPEA2	0.5063	0.3412	16.52	0.0696	0.0603
	awGA	0.5625	0.2767	28.58	0.0572	0.0619

and awWA on C measure. Such better convergence should mainly attribute to the hybrid sampling strategy of VEGA's preference for the edge region of the Pareto front and PDDR-FF's tendency converging toward the center area of the Pareto front. They preserve better performances both in efficacy and efficiency. Especially, h-MoEDA can also keep diversity evenly without special distribution mechanisms like NSGA-II and SPEA2.

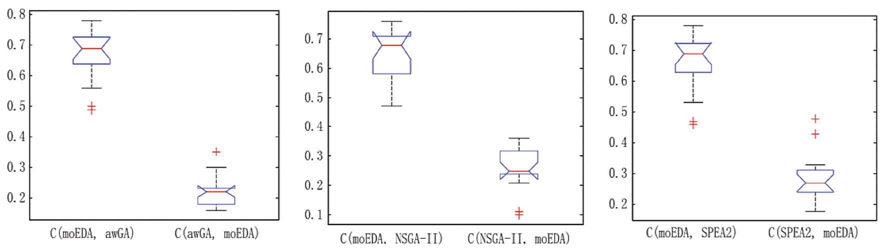


Fig. 6. Box plot of coverage measure conducted on LA35

4.2 Distribution Performance

The experiment results of dispersion performances, SP are shown in Table 4 and Fig. 7 presents the distribution range of space conducted on LA35. It indicates that MoEDA is obviously better than awGA, while better than NSGA-II and SPEA2. Without special mechanism to preserve the diversity evenly, h-MoEDA can also achieve satisfactory dispersion performance.

4.3 Computation Costs

The computation costs of evolutionary-based algorithms mainly depend on the number of fitness evaluations. We need to compare h-MoEDA with other methods, and the termination criterion is to reach the maximum number of iterations

Table 4. Comparison of the spacing measure by h-MoEDA, NSGA-II, and SPEA2 and AWGA

Problem	Algorithm	Mean		Improved	Sd	
		SP (h-MoEDA)	SP(A)	(SP(h-MoEDA) SP(A))/SP(A)(%)	SP(h-MoEDA)	SP(A)
LA29	NSGA-II	16.3568	20.1712	−18.91	4.4081	6.4981
	SPEA2	16.3568	22.2868	−26.62	4.4081	6.7942
	awGA	16.3568	25.8421	−36.70	4.4081	7.2521
LA35	NSGA-II	19.0001	24.8423	−20.31	4.9569	6.2297
	SPEA2	19.0001	24.9510	−23.85	4.9569	7.0526
	awGA	19.0001	33.2410	−42.84	4.9569	7.1686
LA38	NSGA-II	17.1736	21.4569	−19.96	5.2566	6.9774
	SPEA2	17.1736	24.0161	−28.49	5.2566	7.1136
	awGA	17.1736	27.0155	−36.43	5.2566	7.9576

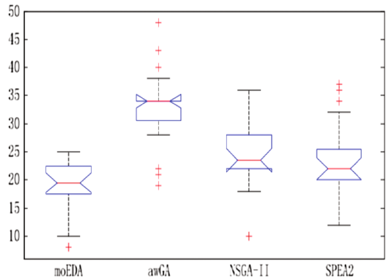


Fig. 7. Box plot of space measure conducted on LA35

(generations). The mean computation cost of h-MoEDA is 217.78 s, NSGA-II and SPEA2 are 245.54, 390.24 s respectively. Compared with SPEA2, SPEA2 needs to calculate all the distances of each individual to all the other individuals first. Then SPEA2 gets the *k-th* nearest distance for each individual after sorting its distances and the time complexity of SPEA2 is $O(mN2 \log N)$ where m is the number of objectives. It is clear that the NSGA-II achieved better computation costs. Without special dispersion preservation mechanism, the time complexity of fitness calculation (PDDR-FF) in h-MoEDA is $O(0)$. It is slightly better than NSGA-II while h-MoEDA needs to estimate the parameters of probability model using the promising solutions.

5 Conclusion

This paper presents an effective h-MoEDA, which solves the MoJSRP. It minimized the expected average makespan and expected total tardiness within a reasonable amount of calculation time. With the framework of the proposed EDA, the explicit probability model of the operation sequence is estimated on the distribution of good solutions found so far and use the constructed model to guide further search behavior. The sampling operator based on the probability model achieves better convergence and stability than the conventional operator such crossover and mutation. In our future work, further experiments will be conducted to determine accuracy of the proposed EDA in response to variations among the parameters.

Acknowledgement. This work is partly supported by the National Natural Science Foundation of China under Grant 61572100, and in part by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS) No. 15K00357.

References

1. Cauty R (2000) Genetic algorithms and engineering optimization. Wiley, New York
2. Deb K, Pratap A, Agarwal S et al (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6(2):182–197
3. Gen M, Cheng R, Lin L (2008) Network models and optimization: multiobjective genetic algorithm approach. Springer, London
4. Giffler B, Thompson GL (1960) Algorithms for solving production-scheduling problems. Oper Res 8(4):487–503
5. Kempf K (1994) Intelligently scheduling semiconductor wafer fabrication. In: Intelligent Scheduling, pp 517–544
6. Michiels W, Aarts E, Korst J (2007) Theoretical aspects of local search. Springer, Berlin
7. Sha DY, Liu CH (2005) Using data mining for due date assignment in a dynamic job shop environment. Int J Adv Manufact Technol 25(11):1164–1174
8. Vieira GE, Herrmann JW, Lin E (2003) Rescheduling manufacturing systems: a framework of strategies, policies, and methods. J Sched 6(1):39–62
9. Vinod V, Sridharan R (2008) Scheduling a dynamic job shop production system with sequence-dependent setups: an experimental study. Rob Comput-Integr Manufact 24(3):435–449
10. Zitzler E, Laumanns M, Thiele L (2001) SPEA 2: improving the strength pareto evolutionary algorithm. In: Evolutionary methods for design, optimization and control. CIMNE, Barcelona

Multi-Queue Priority Based Algorithm for CPU Process Scheduling

Usman Rafi[1], Muhammad Azam Zia[2,3(✉)], Abdul Razzaq[4], Sajid Ali[5], and Muhammad Asim Saleem[6]

[1] Department of CS and IT, UOS Lyallpur Campus, Faisalabad, Pakistan
[2] State Key Laboratory of Networking and Switching Technology, BUPT, Beijing, The People's Republic of China
[3] University of Agriculture, Faisalabad, Pakistan
zia_uaf@hotmail.com
[4] Department, NFC Institute of Engineering and Technology, Multan, Pakistan
[5] Department, D.G. Khan Campus, University of Education, Lahore, Pakistan
[6] Department of IT, Government College University, Faisalabad, Pakistan

Abstract. Operating Systems acts as a base software and acts as a driver for both application programs and system programs. All the programs residing in an operating system has to become process for execution. A modern computer system supports multitasking by single user or multiple users. Different processes have different priorities. The major goal of an operating system is to reduce waiting time and enhance throughput by scheduling processes in some way. This paper discusses various scheduling terms and scheduling algorithms. We have proposed a new approach for scheduling. This proposed algorithm is based on the mixture of MQMS, Priority Scheduling mechanism and Round Robin scheduling. The proposed algorithm facilitates operating system by managing separate queue for separate priority of process and manages queue scheduling in round robin fashion with dynamic time slicing. Processes are added to appropriate queue and this decision is based on any user defined or system defined criteria. We have also discussed various case studies regarding this algorithm and compared its results with priority scheduling algorithm. These case studies are limited to two queuing system up till now. We have also proposed multiple queue management (more than 2), dynamic time slicing instead of half execution scheme and varying execution times of queues as future work of this algorithm scheme.

Keywords: Priority scheduling · Multi-queue scheduling · Dynamic time slice execution · Fair priority scheduling · Single processor multi-queue scheduling · Improved priority scheduling

1 Introduction

A computer system comprises of software and hardware. Software includes programs for system and user needs. Hardware is a set of resources including CPU,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_4

GPU, RAM and many more of them. The most important software is Operating System that acts as an interfacing layer which serves the users in using hardware effectively. User needs are fulfilled by set of Application Software that requires a base software to run i.e. operating system. Software may be in running state or it may not be running. The former is called a Process while the latter is called Program. A process requires certain resources for execution. CPU and RAM are major resources that each process requires for execution. CPU is the most costly and important resource. It must be utilized properly. CPU should never be left underutilized. Operating System must ensure proper utilization of the CPU.

A computer system may be Single Processor System or it may be Multi-Processor System. Single processor system allows only one process to acquire the CPU and gets execute on it. On the other hand multi-processor system allows more than one processes to be executed with one on each processor at a time.

A process changes its activity from time to time as it executes. Process activity is defined in terms of two-state model or five-state model. Two-state model defines that a process can have two states: running and not running. It gives an abstract idea. Two state model is shown in Fig. 1. Stallings [21] described a more elaborative version of two state model called five-state model. Five state model took into account the reason of process for not being in the running state. So, five-state model described five process activities: new, ready, running, waiting and terminated. Figure 2 shows five state model for process execution states.

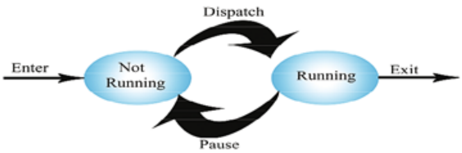


Fig. 1. Two state model for process states

Fig. 2. Five state model for process states

Section 2 of the paper describes CPU Scheduling, its need, various queue management systems, queues involved in scheduling and scheduling criteria. Section 3 is organized to give a brief idea of various scheduling algorithms. Section 4 describes some of related work. Sections 5 and 6 describes our proposed algorithm, its working logic and pseudo code. Section 7 comprises of various case studies and results of our proposed algorithm. The results of our proposed algorithm are compared with the priority scheduling algorithm in Sect. 8.

2 CPU Scheduling

Operating system may be designed to be Single Programmed Systems or Multi Programmed Systems.

A single programmed system can only accommodate single process in memory for execution. CPU has to remain idle in the case of process waiting for some I/O operation. This results in underutilization and causes increase in waiting time and turnaround time.

Silberchatz et al. [18] narrated that a multi-programmed system facilitated the efficient utilization of processor by allowing multiple processes or jobs ready/waiting (residing in memory) for execution and switching the CPU among these processes. If an executing process had to wait for some I/O operation, the CPU no longer remained idle. It switched to the next job waiting for CPU allocation to get executed. In this manner proper utilization of CPU was achieved. To achieve this utilization, scheduling was required. Scheduling scheme comprised of criteria and algorithm to implement this criteria. CPU scheduling was of immense importance in multi-programmed systems because CPU was expansive and scarce resource.

2.1 Ready Queue Management

The ready queue is a bulk of processes ready to get executed. Ready queue may be partitioned as it is done in Multilevel Queue Scheduling and Multilevel Feedback Scheduling. There are following two ways to manage ready queue:

(1) SQMS
SQMS (Single Queue Management System) involves no partitioning of ready queue. SQMS consists of single partition ready queue in which all the incoming processes are stored. It is a straight forward queue management scheme. It is due to the fact that processes are to be picked for execution from a single queue. So operating system requires only process scheduling algorithm. Only a single queue has to be managed and scheduled. No overhead is involved in terms of scheduling ready queue partitions or multiple ready queues. In addition to this, operating system doesn't require any overhead for deciding the addition of processes to un-partitioned ready queue.

(2) MQMS
MQMS (Multi-Queue Management System) scheme maintains multiple ready queues of jobs or partitions of ready queue. Processes are added to appropriate ready queue portion depending upon certain criteria. MQMS scheme is complex scheme in terms of implementation, management and scheduling. The selection of process for execution requires two important decisions: (i) appropriate queue selection and (ii) appropriate process selection from the selected queue. Hence MQMS involves both queue scheduling algorithm and process scheduling algorithm. MQMS scheme involves overhead in terms of queue selection.

A process is first added to job queue. After this, it is moved to ready queue. Process can now move itself among ready queue, running state and waiting queue as and when required. All these state changes depend upon a criteria on the basis of which the Operating System decides that which process move to ready queue, which started execution next, which process will be needed to move to waiting queue and which process pre-empted and moved to ready queue. These decisions are taken by intelligent algorithms, executed by CPU Scheduler.

3 Scheduling Algorithms

Tanenbaum [22] described that First Come First Served (FCFS), Shortest Job First (SJF), Shortest Remaining Time First (SRTF), Round Robin (RR) and Priority Scheduling (PS) algorithms were commonly used Scheduling Algorithms. These algorithms were implemented in the form of SQMS. However, priority scheduling may had MQMS version. Mishra and Khan [11] narrated that SJF only permitted process with lowest burst time to acquire CPU for execution and lowered average waiting time. RR scheduling defined a time quantum for assigning CPU fairly to each process present in the ready queue for specific amount of time. We discuss Shortest Job First (SJF), Shortest Remaining Time First (SRTF), Round Robin (RR) and Priority Scheduling algorithms below.

3.1 Shortest Job First (SJF)

Shortest Job First is based on the Burst Time of processes. SJF is based on a simple idea derived from FCFS, that is, executing smaller processes before larger ones lowers the Average Waiting Time (AWT) of process set. If a sub set of processes have same burst time, then they are executed in FCFS fashion. The processes are arranged in a queue on the base of their Burst Times in a way that process with smallest burst time is placed at the Head/Front of queue while the process with largest burst time is placed at Rear of queue. In this way, CPU Scheduler always picks the process with smallest burst time from ready queue. It's a non-preemptive scheduling algorithm. The main advantage of SJF is to decrease the Average Waiting Time of process set. However, this algorithm does not take into account process priorities. SJF may result in starvation of larger processes.

3.2 Shortest Remaining Time First (SRTF)

Shortest Remaining Time First scheduling algorithm is simply preemptive version of SJF. In SRTF, the main idea is simply the same, that is, execute small process before large one. But in this algorithm, if a process is under execution and during course of its execution a new process enters the system. Then the burst time of newly entered process is compared with remaining burst time of currently executing process. If the comparison results in the fact that entering process has smaller burst time, then CPU is preempted and allocated to the newly entering process. However, if the new entering process has same burst time as that of already executing process, then FCFS manner is used for their scheduling. The major advantage is the exploitation of phenomenon to lower Average Waiting Time of process set. The demerit of SRTF is that, the processes in real environment have priorities and SRTF does not support priorities. Moreover, it is not necessary that smaller processes are important while larger ones are less important. This algorithm also includes high overheads for queue managements, queue rearrangements, burst time comparisons and context switches. It may also give rise to starvation for large processes.

3.3 Round Robin (RR)

Round Robin scheduling algorithm is same as that of FCFS with certain innovation. This innovation includes preemption and Time Quantum. Time Quantum is a small amount of time. Processes are arranged in a circular queue. CPU is allocated a process from the queue; CPU executes the allocated process for a time period equal to that of time quantum. When the time quantum expires, the CPU is preempted and next process from the queue is selected. Next process is also executed for a time length equal to time quantum. This algorithm continues to schedule processes in this fashion until queue is empty. The major advantage of this scheme is that each of the process gets fair share and equal opportunity of execution. The major drawback is that, if time quantum is very small, then too many context switches are involved. On the other hand, if time quantum is large, then RR may perform same as FCFS. Moreover, RR also does not consider priorities of processes. Siregar [23] narrated that starvation could be avoided if CPU was allocated fairly to processes. Time quantum was an important factor in RR scheduling scheme for lowering waiting time of process set. Time quantum could take any value. But best time quantum should be discovered for minimizing average waiting time.

3.4 Priority Scheduling (PS)

Priority Scheduling s involves assigning a Priority to each process. A Priority is either indicated by an integer number (0, 1, 2, 3, etc.) or a string (low, medium, high etc.). In the case of numeric priority, two choices exist: higher number may indicate a higher priority (our proposed algorithm consideration) or lower priority (other algorithm consideration). This scheduling algorithm is implemented using a Priority Queue. Processes are arranged in decreasing order of priority (Front to Rear) with process having highest priority placed at front while process having second highest priority placed at a place next to front and so on. The CPU Scheduler picks process with highest priority and allocates CPU for execution. If a number of processes have same priority then they are scheduled in FCFS fashion as of in SJF and RR case. PS may be preemptive or non-preemptive. If non-preemptive version is being used, then no matter how much higher priority process enters the queue, if CPU is executing a process then it will not be preempted unless the process is completed. On the other hand, the preemptive version preempted the CPU if such a situation occurs. The main attraction in PS is realistic approach about priorities. However, PS can lead to starvation situation preventing lower priority processes from execution. Its implementation involves high overheads of queue rearrangements and priority comparisons in case of preemptive version. PS can be implemented in the form of either SQMS (Traditional Priority Scheduling) or MQMS (Multi-Level Queue Scheduling). Rajput and Gupta [14] narrated that in priority scheduling, processes were arranged in queue according to their fixed priority values. A high priority incoming process interrupted low priority process. Waiting time of low priority processes were thus increased and hence starvation occurred.

Adekunle et al. [2] analysed the performance of FCFS, SJF, RR, SRTF and Lottery scheduling algorithms. SJF showed higher while RR showed medium CPU utilization. SJF also showed lower average waiting time and good response time as compared to RR. However, RR algorithm was found to be fair and starvation free.

Almakdi et al. [4] simulated FCFS, SJF, RR and SRT algorithms for analysing scheduling algorithm performance. The results showed that the lowest average waiting time and the lowest average turnaround time were showed by SJF. However, high average core utilization and high average throughput were resulted by RR.

4 Related Work

Rao et al. [15] elaborated that Multilevel Feedback Queue Scheduling involved use of multiple queues. Process could be migrated (promoted/demoted) from one queue to another. Each queue had its own scheduling algorithm. Similarly, criteria were also defined for promotion and demotion of processes. PMLFQ resulted in starvation because of execution of high priority process queues before low priority process queues.

Goel and Garg [5] explained fairness was an important scheduling criteria. Fairness ensured that each process got fair share of CPU for execution. Further, higher priority processes exhibited smaller waiting time and starvation could take place for lower priority process.

Shrivastav et al. [16] proposed Fair Priority Round Robin with Dynamic Time Quantum (FPRRDQ). FPRRDQ calculated time quantum on the basis of priority and burst time for each individual process. Experimental results when compared with Priority Based Round Robin (PBRR) and Shortest Execution First Dynamic Round Robin (SEFDRR) scheduling algorithms, showed that FPRRDQ had improved performance considerably.

Patel and Patel [13] proposed a Shortest Job Round Robin (SJRR) scheduling algorithm. SJRR suggested to arrange processes in SJF fashion. The shortest process's burst time was designated as time quantum. Case study showed reduction in average waiting time of process set.

Abdulrahim et al. [1] proposed New Improved Round Robin (NIRR) algorithm for process scheduling. The algorithm involved two queues ARRIVE and REQUEST. In this algorithm, the time quantum was calculated by finding the ceiling of average burst time of process set. First process in the REQUEST queue was allocated to CPU for one time quantum. In any case, the executing process had burst time less than or equal to half time quantum then CPU reallocated for remaining CPU burst time. Otherwise, process was removed from REQUEST and added to ARRIVE. NIRR algorithm improved scheduling performance and also lowered number of context switches as the case of RR algorithm.

Mishra and Rashid [12] introduced Improved Round Robin with Varying Time Quantum (IRRVQ) scheduling algorithm. IRRVQ proposed arrangement of processes (in ready queue) in SJF fashion and setting time quantum equal to burst time of first process. After a complete execution of first process, next processes were selected and assigned one time quantum. If finished, processes were removed from ready queue otherwise, processes were preempted and added to tail of ready queue.

Sirohi et al. [20] described an improvised round robin scheduling algorithm for CPU scheduling. This algorithm calculated the time quantum by finding the average of the burst times of processes. The processes were allocated to ready queue in SJF order. CPU was allocated to first process for one time quantum. The reallocation of CPU was based on remaining burst time of process under execution.

Akhtar et al. [3] proposed a preemptive hybrid scheduling algorithm based on SJF. Initially, it showed similar steps as that of SJF. The algorithm reduced number of context switches considerably but only showed minor change in average waiting time. The algorithm could exhibit starvation.

Joshi and Tyagi [7] explained Smart Optimized Round Robin (SORR) scheduling algorithm for enhancing performance. At first step, the SORR algorithm arranged processes in increasing order of burst times. In the second step, the mean burst time was calculated. Third step involved calculation of Smart Time Quantum (STQ) after which the CPU was allocated to first process in the ready queue for a time quantum equal to calculated STQ. If the allocated process was remained with burst time less or equal to 1 time quantum, CPU was reallocated for the burst time. Otherwise process was preempted, next process was selected from ready queue and CPU was allocated to the newly selected process for STQ.

Lulla et al. [10] devised a new algorithm for CPU scheduling. The algorithm introduced mathematical model for calculating initial time quantum and dynamic time quantum. Dynamic time quantum assisted in lowering number of context switches. The algorithm proposed execution of top two high priority processes by assigning calculated initial time quantum followed by the calculation and assignment of dynamic time quantum for execution of top two high priority processes.

Goel and Garg [6] proposed Optimized Multilevel Dynamic Round Robin Scheduling (OMDRRS) algorithm. This algorithm suggested the allocation of dynamic time slices. Factor analysis based on arrival time, burst time and priority was calculated for sorting processes in ready queue. Burst values of first and last process were used for time slice calculation. Simulated results, statistically analysed using ANOVA and t-test, showed considerable performance improvement.

Kathuria et al. [8] devised Revamped Mean Round Robin (RMRR) scheduling algorithm. RMRR involved two queues ReadyQueue (RQ) and Pre-ReadyQueue (PRQ). Time quantum was calculated by taking mean of the burst time of processes. The proposed algorithm lowered average waiting time, average turnaround time and number of context switches as compared to RR scheduling. RMRR showed the best response time when compared with FCFS, SJF and RR scheduling algorithms.

Singh and Singh [19] proposed a hybrid scheduling algorithm based on Multi Level Feedback Queue (MLFQ) and Improved Round Robin (IRR) scheduling principles. The proposed algorithm suggested the sub-division of memory in three queues as in MLFQ. Fonseca and Fleming's Genetic Algorithm (FFGA) was implemented for optimization. Results showed optimized performance in comparison to various algorithms (Fig. 3).

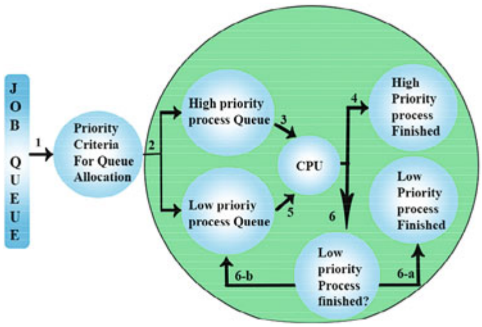


Fig. 3. Schedulers and scheduling queues

5 Proposed Algorithm

In this paper, we proposed a new scheduling algorithm as an extension of priority scheduling. Here, we proposed it only for single processor system (neither multiple processors nor multi-core processors). It is based on hybrid concepts of multiple queue management, priority scheduling and round robin scheme. In our algorithm, multiple queue management scheme facilitates subdivision of ready queue in multiple portions. Shukla et al. [17] described the use of multi-level queuing approach. Multi-level queuing involved partitioning of ready queue into various sub queues. In such scheduling, processes were assigned permanently to specific queue. Memory size, process type and priority were some of parameters used for such assignments.

In our proposed algorithm, each job is allocated to specific queue according to criteria predefined criteria. The criteria consists of two factors: (1) Priority of individual process and (2) Burst Time of process. Firstly, a priority criteria is defined in a set of processes. The processes are allocated after comparing process priorities to priority criteria. Our proposed algorithm assumes that the largest number represented highest priority and smallest number represented lowest priority. Shortest job first scheme is applied while arranging processes when they are transferred from the job pool to the ready queue partitions. Jobs inside a specific queue are scheduled as defined by priority scheduling scheme. Khan and Kakhani [9] suggested SJF based PS was proposed for performance improvement. Processes were placed in ready queue in higher to lower order of priority. SJF was employed for arranging process in case of more than one processes having similar priority. Results of SJF based PS were compared to FCFS based PS and results showed considerable decrease in average waiting and average turnaround times. Moreover, the queues are scheduled in a round robin fashion from the highest priority queue to the lowest priority queue. Round robin scheme is based on dynamic time slicing. Whenever a queue is given time slice it is according to the process burst time and some variations of it. Our algorithm has following main steps:

Step 1. Create Multiple Queues (one for high priority processes and other for low priority processes) based on Priority Scheme. These queues are designated as High Priority and Low Priority Queues.

Step 2. Create Processes in Job Pool.

Step 3. Define Priority Criteria.

Step 4. Allocate processes to queues based on priority criteria and burst time. Processes are arranged in decreasing order of priority from front to rear. Additionally following are considered:

- **Step 4.1.1** If process priority is less than or equal to priority criteria then add it to the Low Priority Queue. Otherwise add it to High Priority Queue.
- **Step 4.1.2** If multiple processes have same priority then arranges them in a way that the process with shortest burst time comes first in the queue.

Step 5. Select first process from High Priority Queue and execute it fully (for entire burst time by assigning time slice equal to the burst time) and remove it from queue head.

Step 6. Select first process from Low Priority Queue and execute for a time equal to half that of burst time (assign time slice equal to half the burst time). Important thing to consider is that, in this case, each low priority process completed in two halves.

Step 7. Repeat Steps 5–6 until both queues are empty.

One of the interesting things to note is that, CPU is alternated between high priority and low priority processes in order to increase the fairness. Our approach lowers the starvation faced by low priority processes in the presence of high priority processes. Another point may arise that why low priority processes are executed in two halves? The answer of this question is that, it is done so that if a low priority process is executed fully, then the purpose of two queues is unclear. Similarly, if high priority process is executed for half time, then the use of priorities becomes meaningless.

6 Pseudo Code

Proposed Algorithm (P[1, $\cdots$, n], Priority Criteria) (Fig. 4)

```
begin
Create Queue Portions QH[1....nh], QL[1....nl]
For each P[i] in P[1.......n]
        If (P[i].Priority <= PriorityCriteria) Then
                Add P[i] to QL in such a way that processes are arranged in
                decreasing order of priority and increasing order of burst time
        Else
                Add P[i] to QH in such a way that processes are arranged in
                decreasing order of priority and increasing order of burst time
              End If
Repeat following steps while QH and QL are not empty
        Select a process from QH head, allocate CPU, execute completely and
        remove it from QH
        Select a process from QL head, allocate CPU, execute for the time half to
        its original burst time
                If (P[i].RBT) = 0 Then
                    Removes it from QL
                Else
                    place it back to QL head
            End If
                Goto step 4
            End while
Completion of process set
End Algorithm
*RBT = Remaining Burst Time
```

Fig. 4. The proposed algorithm

7 Case Studies

7.1 Case Study 1

Gantt Chart of above process data when executed by Priority Scheduling is as Fig. 5 (Table 1).

Table 1. Process set with five processes

Process	Burst time	Priority
P1	10	3
P2	1	1
P3	2	3
P4	1	4
P5	5	2

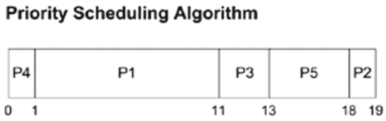


Fig. 5. Gantt chart for priority scheduling algorithm

Average Waiting Time = 8.6 ms.
Average Turnaround Time = 12.4 ms.
Now we define certain parameters related to proposed algorithm.
(1) Priority Criteria: Process with Priority > 2 is placed in High Priority Queue while process with Priority <= 2 is placed in Low Priority Queue.
(2) Ready Queue Partitions: As Fig. 6.
The proposed algorithm on given data results in Fig. 7.

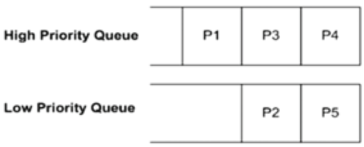


Fig. 6. Scheduling queues

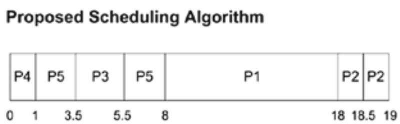


Fig. 7. Gantt chart for proposed algorithm

Average Waiting Time = 6.5 ms
Average Turnaround Time = 10.3 ms

7.2 Case Study 2

Gantt Chart of above process data when executed by Priority Scheduling is as Fig. 8 (Table 2).

Average Waiting Time = 7.0 ms.
Average Turnaround Time = 10.8 ms.
Now we define certain parameters related to proposed algorithm.

Table 2. Process set with five processes

Process	Burst time	Priority
P1	10	3
P2	1	1
P3	2	4
P4	1	5
P5	5	2

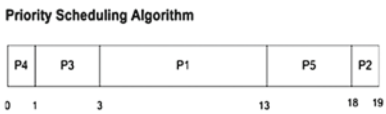


Fig. 8. Gantt chart for priority scheduling algorithm

(1) Priority Criteria: Process with Priority $>$ 2 is placed in High Priority Queue while process with Priority $\leq$ 2 is placed in Low Priority Queue.
(2) Ready Queue Partitions: As Fig. 9.
The proposed algorithm on given data results in Fig. 10.

Average Waiting Time = 6.5 ms.
Average Turnaround Time = 10.3 ms.

7.3 Case Study 3

Gantt Chart of above process data when executed by Priority Scheduling is as Fig. 11 (Table 3).

Average Waiting Time = 33.8 ms.
Average Turnaround Time = 49.8 ms.
Now we define certain parameters related to proposed algorithm.
(1) Priority Criteria: Process with Priority $>$ 2 is placed in High Priority Queue while process with Priority $\leq$ 2 is placed in Low Priority Queue.
(2) Ready Queue Partitions: As Fig. 12
The proposed algorithm on given data results in Fig. 13.

Average Waiting Time = 33.5 ms
Average Turnaround Time = 49.5 ms

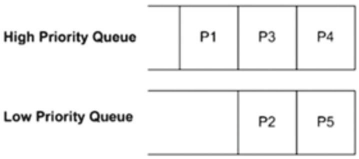


Fig. 9. Scheduling queues

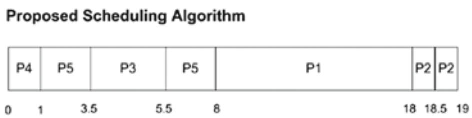


Fig. 10. Gantt chart for proposed algorithm

Table 3. Process set with five processes

Process	Burst time	Priority
P1	12	3
P2	19	3
P3	21	5
P4	13	2
P5	15	3

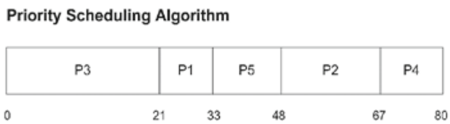


Fig. 11. Gantt chart for priority scheduling algorithm

7.4 Case Study 4–Special Case

This case deals with a situation involving same priority processes (high priority). Similarly, inverse case of this can also be assumed. In such cases all the processes added to a single queue either high or low priority (Table 4).

Gantt Chart of above process data when executed by Priority Scheduling is as Fig. 14.

Average Waiting Time = 3.67 ms
Average Turnaround Time = 9.33 ms
Now we define certain parameters related to proposed algorithm.
(1) Priority Criteria: Process with Priority > 2 is placed in High Priority Queue while process with Priority ≤ 2 is placed in Low Priority Queue.
(2) Ready Queue Partitions: As Fig. 15
The proposed algorithm on given data results in Fig. 16.

Average Waiting Time = 3.67 ms
Average Turnaround Time = 9.33 ms

In the case with all the processes getting place in a single queue, the proposed algorithm gives an average waiting time and average turnaround time exactly equal to priority scheduling algorithm.

Fig. 12. Scheduling queues

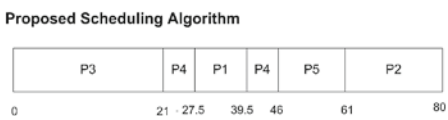


Fig. 13. Gantt chart for proposed algorithm

Table 4. Process set with three processes

Process	Burst Time	Priority
P1	10	5
P2	3	6
P3	4	7

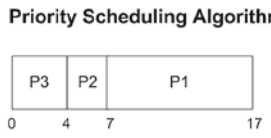


Fig. 14. Gantt chart for proposed scheduling algorithm

8 Results and Discussion

8.1 Comparison of Average Waiting Time

In this section we compare average waiting time of our proposed algorithm and traditional priority scheduling algorithm. This comparison is based on case studies discussed in the previous sections.

The average waiting time is minimized by our proposed algorithm as compared to the baseline scheduling algorithm as shown in Fig. 17. However, the last case showed same average waiting time. This was because, the last case comprised of special dataset involving all the processes with same priority (high). In this case, only one priority queue was formed rather two queues. So, in this case it exhibited same behaviour as that of traditional priority scheduling algorithm.

8.2 Comparison of Average Turnaround Time

In this section we presented comparison of average turnaround time of our proposed algorithm with traditional priority scheduling algorithm. This comparison based on case studies discussed in the previous sections.

Figure 18 clearly showed that in first three case studies the average turnaround time of our proposed scheduling algorithm was less than traditional priority scheduling algorithm. However, the last case showed same average turnaround time. This was because, the last case comprised of special dataset having

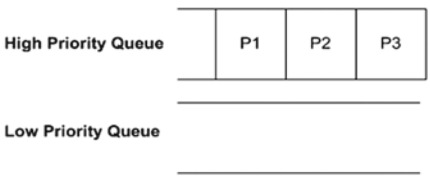


Fig. 15. Scheduling queues

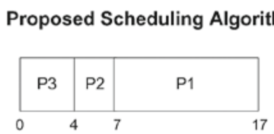


Fig. 16. Gantt chart for proposed algorithm

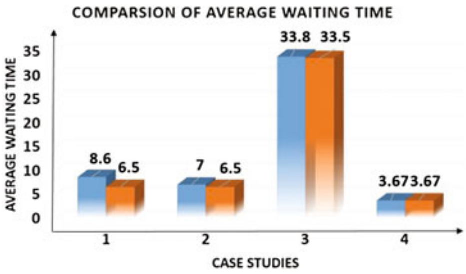


Fig. 17. Comparison of average waiting time

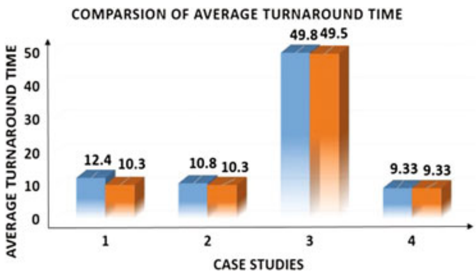


Fig. 18. Comparison of average turnaround time

all the processes having high priority. In this case only one priority queue was formed rather two queues.

9 Conclusion and Future Work

It is concluded that our proposed algorithm used minimum waiting and turnaround time as compared to baseline Priority Scheduling algorithms. Our proposed algorithm also supported multiple queue supporting operating systems using different priorities. Each type of priority can be managed with a separate queue. The proposed algorithm provided fairness to both high priority and low priority jobs and thus saved low priority jobs from starvation. The importance is given to high priority process within both high priority and low priority list by our proposed algorithm. Our algorithm also enhanced throughput by lowering average waiting and average turnaround time. As compared to Round Robin scheduling with small time slice, proposed algorithm used a scheme requiring only two context switches per process. In this way, our proposed algorithm limited the number of context switches required for a job. This algorithm is a good mixture of priority scheduling, shortest job first and round robin scheduling.

As a future work, we will try out different execution lengths or time slices for high priority processes and low priority processes to find optimized solution providing optimized results of scheduling criteria. We will also extend this algorithm in a way that for each priority type a single queue may be used.

References

1. Abdulrahim A, Abdullahi SE, Sahalu JB (2014) A new improved round robin (nirr) cpu scheduling algorithm. Int J Comput Appl 90(4):27–33
2. Adekunle O (2014) A comparative study of scheduling algorithms for multiprogramming in real-time systems. Int J Innov Sci Res 12:180–185
3. Akhtar M, Hamid B et al (2015) An optimized shortest job first scheduling algorithm for cpu scheduling. J Appl Environ Biol Sci 5:42–46
4. Almakdi S (2015) Simulation and Performance Evaluation of CPU Scheduling Algorithms. LAP LAMBERT Academic Publishing, Saarbrücken
5. Goel N, Garg RB (2013) A comparative study of cpu scheduling algorithms. Int J Graph Image Process 2(4):245–251
6. Goel N, Garg RB (2016) Performance analysis of cpu scheduling algorithms with novel omdrrs algorithm. Int J Adv Comput Sci Appl 7(1):216–221
7. Joshi R, Tyagi SB (2015) Smart optimized round robin (sorr) cpu scheduling algorithm. Int J Adv Res Comput Sci Softw Eng 5:568–574
8. Kathuria S, Singh PP et al (2016) A revamped mean round robin (rmrr) cpu scheduling algorithm. Int J Innov Res Comput Commun Eng 4:6684–6691
9. Khan R, Kakhani G (2015) Analysis of priority scheduling algorithm on the basis of fcfs and sjf for similar priority jobs. Int J Comput Sci Mob Comput 4:324–331
10. Lulla D, Tayade J, Mankar V (2015) Priority based round robin cpu scheduling using dynamic time quantum. Int J Emerg Trends Technol 2:358–363
11. Mishra MK (2012) An improved round robin cpu scheduling algorithm. J Glob Res Comput Sci 3(6):64–69
12. Mishra MK, Rashid F (2014) An improved round robin cpu scheduling algorithm with varying time quantum. Int J Comput Sci Eng Appl 4(4):1–8
13. Patel R, Patel M (2013) Sjrr cpu scheduling algorithm. Int J Eng Comput Sci 2:3396–3399
14. Rajput G (2012) A priority based round robin cpu scheduling algorithm for real time systems. Int J Innov Eng Technol 1:1–10
15. Rao MVP, Shet KC, Roopa K (2009) A simplified study of scheduler for real time and embedded system domain. Comput Sci Telecommun 12(5):1–6
16. Shrivastav MK, Pandey S et al (2012) Fair priority round robin with dynamic time quantum: Fprrdq. Int J Mod Eng Res 2:876–881
17. Shukla D, Ojha S, Jain S (2010) Data model approach and markov chain based analysis of multi-level queue scheduling. J Appl Comput Sci Math 8(4):50–56
18. Silberschatz A, Gagne G, Galvin PB (1983) Operating System Concepts, 8th edn. Addison-Wesley Pub. Co, Boston Binder Ready Version
19. Singh N, Singh Y (2016) A practical approach on mlq-fuzzy logic in cpu scheduling. Int J Res Educ Sci Methods 4:50–60
20. Sirohi A, Pratap A, Aggarwal M (2014) Improvised round robin (cpu) scheduling algorithm. Int J Comput Appl 99(18):40–43
21. Stallings W (2011) Operating Systems–Internals and Design Principles, 7th edn. DBLP
22. Tanenbaum AS, Tanenbaum AS (2001) Modern Operating Systems, 2nd edn. Prentice-Hall, Upper Saddle River
23. Ulfahsiregar M (2012) A new approach to cpu scheduling algorithm: Genetic round robin. Int J Comput Appl 47(19):18–25

An Order-Based GA for Robot-Based Assembly Line Balancing Problem

Lin Lin[1,2,3](✉), Chenglin Yao[1], and Xinchang Hao[4]

[1] Dalian University of Technology, Dalian 116024, People's Republic of China
lin@dlut.edu.cn
[2] Fuzzy Logic Systems Institute, Tokyo, Japan
[3] Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian 116024, People's Republic of China
[4] Waseda University, Tokyo 169-0051, Japan

Abstract. In the real world, there are a lot of scenes from which the product is made by using the robot, which needs different assembly times to perform a given task, because of its capabilities and specialization. For a robotic assembly line balancing (rALB) problem, a set of tasks have to be assigned to stations, and each station needs to select one robot to process the assigned tasks. In this paper, we propose a hybrid genetic algorithm (hGA) based on an order encoding method for solving rALB problem. In the hGA, we use new representation method. Advanced genetic operators adapted to the specific chromosome structure and the characteristics of the rALB problem are used. In order to strengthen the search ability, a local search procedure is integrated under the framework the genetic algorithm. Some practical test instances demonstrate the effectiveness and efficiency of the proposed algorithm.

Keywords: Balancing · Improvement · Genetic algorithms · Local search · Neighborhood structure · Robotic assembly line

1 Introduction

An assembly line system is a manufacturing process in which interchangeable parts are added to a product in a sequential manner to produce a finished product. It is important that an assembly line is designed and balanced so that it works as efficiently as possible. Most of the work related to the assembly lines concentrate on the assembly line balancing (ALB). The ALB model deals with the allocation of the tasks among stations so that the precedence relations are not violated and a given objective function is optimized.

Since the ALB model was first formulated by Helgeson et al. [13], many versions of ALB arise by varying the objective function [26]: Type-F is an objective independent problem which is to establish whether or not a feasible line balance exists. Type-1 and Type-2 have a dual relationship; the first one tries to minimize the number of stations for a given cycle time, and the second one tries

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_5

to minimize the cycle time for a given number of stations. Type-E is the most general problem version, which tries to maximize the line efficiency by simultaneously minimizing the cycle time and a number of stations. Finally, Type-3, 4 and 5 correspond to maximization of workload smoothness, maximization of work relatedness and multiple objectives with Type-3 and Type-4, respectively [15]. Furthermore, most versions of above models are NP-hard [12].

Recently, genetic algorithm (GA) and other evolutionary algorithms (EAs) have been successfully applied in a wide variety of ALB problems. Falkenauer and Delchambre [6] were the first to solve ALB with GAs. Following application of GAs for solving ALB model was studied by many researchers, e.g., Anderson and Ferris [2], Rubinovitz and Levitin [22], Gen et al. [10], Bautista et al. [3], Sabuncuoglu et al. [24], Goncalves and Almeida [11], Brown and Sumichrast [4], and Nearchou [21]. However, most of the researchers focused on the simplest version of the problem, with single objective and ignored the recent trends, i.e., mixed-model production, u-shaped lines, robotic lines and etc.

Sigeru focused on implementing particle swarm optimization (PSO) to optimize the robotic assembly line balancing (RALB) problems with an objective of maximizing line efficiency. By maximizing the line efficiency, industries tend to utilize their resources in an efficient manner [14]. Slim studied a robotic assembly line balancing problem which goal is to maximize the efficiency of the line and proposed some resolution methods which define the suitable component and point positions in order to define the strategy of pick and place for each robot [5]. In Nilakantan's research, bio-inspired search algorithms, viz. particle swarm optimization (PSO) algorithm and a hybrid cuckoo search and particle swarm optimization (CS-PSO) were proposed to balance the robotic assembly line with the objective of minimizing the cycle time [20]. Armin et al. extended the existing algorithms to solve the robotic assembly line problem [27]. Yoosefelahi et al. [29] considered a different type II robotic assembly line balancing problem (RALB-II). One of the two main differences with the existing literature is objective function which is a multi-objective one. The aim is to minimize the cycle time, robot setup costs and robot costs. The second difference is on the procedure proposed to solve the problem. In addition, a new mixed-integer linear programming model is developed. Snke [16] introduced a new robot concept that aims at closing the gap between a manual assembly and a fully automatic assembly. It is intended to be used for handling and assembly of small parts in a highly agile production scenario, which employs both human workers and robots in the same line, with a frequent need for reconfiguration. The objective is to minimize the cost of the stations. This formulation is very similar to the rALB problem.

Gao and Gen [7] proposed a hybrid GA for solving rALB-2, in which different robots may be assigned to the assembly line tasks, and each robot needs different assembly times to perform a given task due to its capabilities and specialization. They focused on the genetic operators adapted to the specific chromosome structure and the characteristics of the rALB problems.

In this paper, we also consider a hybrid GA approach for solving rALB-2 problems. Different with prophase works [7], we not only focus on the design of GA to minimize the cycle time of the assembly systems, but also hammer at the key point of improvement of assembly system depended on the final result analysis. We propose a hybrid genetic algorithm (hGA) to solve the rALB-2. The genetic algorithm uses partial representation technique, that is, the coding space only contains a part of the feasible solutions, in which the optimal solution is included. New crossover and mutation operators are also developed to adapt to the specific chromosome structure and the nature of the problem. In order to strengthen the search ability, a local search procedure works under the framework of GA, this only investigates the neighbor solutions that have possibilities to improve the solutions of current population.

The rest of the paper is organized as follows. In Sect. 2, we formulate the mathematical model of rALB-2. The proposed GA for the problem is described in Sect. 3. In Sect. 4, introduce a local search procedure. In Sect. 5, we demonstrate effectiveness comparing with different GA approaches. This paper gives the conclusion in Sect. 6.

2 Robotic Assembly Line Balancing Problem

The assembly of each product unit requires the execution of n tasks (indivisible elements of work). Precedence constraints partially specify the order in which the tasks have to be performed. They can be represented by an acyclic precedence graph which contains nodes for all tasks. As mentioned above, the type II robotic assembly line balancing problem usually occurs when changes in the production process of a product take place. In this case, the robotic assembly line has to be reconfigured using the present resources (such as robots) so as to improve it efficiency for the new production process. The problem concerns how to assign the tasks to stations and how to allocate the available robots for each station in order to minimize the cycle time under the constraint of precedence relationships.

Let assembly tasks ($i = 1, 2, \cdots, 10$) assigned on 4 stations. Robots (R1, R2, R3, R4) are to be equipped on the 4 stations. The balancing chart of the solution was drawn for analyzing the solution. Figure 1 shows the idle time of the station 1, 2, 3 is very big, and it also means this line did not get the balancing for producing. In the real world, the assembly line is not just for producing one unit of the product. It should produce several units. We give the Gantt chart about 3 units for analyzing the solution just like Fig. 2.

We can see the waiting time occurs in the Fig. 2, and the waiting time means the idle time in the line. For example, the part waiting is occurred when the maximum processing time of the station, which is before the current station, is larger than the processing time of the current station. It means the current station will wait for the parts to process, which come from the anterior station. The processing waiting is occurred when the processing time of the current station is smaller than the processing time of the next station. It means the parts which were produced by the current station will wait for being processed

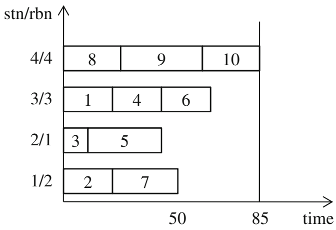


Fig. 1. Balancing chart of the feasible solution

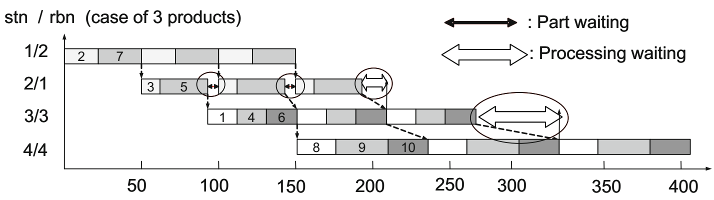


Fig. 2. Gantt chart for producing 3 units

by the next station. Both of the waiting time is the idle time of the assembly line. We want to reduce it by getting the balance of the line. The notation used in this section can be summarized as follows:

Indices:

i, j: index of assembly tasks, $i, j = 1, 2, \cdots, n$;
k: index of stations, $k = 1, 2, \cdots, m$;
l: index of robots, $l = 1, 2, \cdots, m$.

Parameters:

n: total number of assembly tasks;
m: total number of stations (robots);
t_{il}: processing time of the $i-th$ task by robot l;
pre(i): the set of predecessor of task i in the precedence diagram.

Decision variables:

$$x_{jk} = \begin{cases} 1, \text{ if task } j \text{ is assigned to work station } k \\ 0, \text{ otherwise;} \end{cases}$$

$$y_{jk} = \begin{cases} 1, \text{ if robot } j \text{ is assigned to work station } k \\ 0, \text{ otherwise.} \end{cases}$$

Problem Formulation:

$$\min\ C_T = \min_{1 \le k \le m} \sum_{i=1}^{n} \sum_{l=1}^{m} t_{il} x_{ik} y_{kl} \tag{1}$$

$$\text{s. t.} \sum_{k=1}^{m} kx_j k - \sum_{k=1}^{m} kx_{ik} \geq 0, \quad \forall j;\ i \in \text{pre}(i) \tag{2}$$

$$\sum_{k=1}^{m} x_{ik} = 1, \forall i \tag{3}$$

$$\sum_{l=1}^{m} x_{kl} = 1, \forall k \tag{4}$$

$$\sum_{k=1}^{m} x_{kl} = 1, \forall l \tag{5}$$

$$x_{ik} \in 0, 1, \forall k, i \tag{6}$$

$$x_{ik} \in 0, 1, \forall l, k. \tag{7}$$

The objective (1) is to minimize the cycle time (C_T). Inequity (2) represents the precedence constraints. It ensures that for each pair of assembly activities, the precedent cannot be assigned to a station behind the station of the successor, if there is precedence between the two activities. Equation (3) ensures that each task has to be assigned to one station. Equation (4) ensures that each station is equipped with one robot. Equation (5) ensures that each robot can only be assigned to one station.

3 Order-Based GA Design

Genetic algorithms (GAs) are powerful and broadly applicable stochastic search and optimization techniques based on principles from evolutionary theory [9]. GA's have been applied to solve various assembly line balancing problems [8,23,28]. We divide this algorithm into 3 phases. It is as follows:

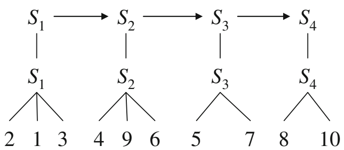


Fig. 3. Structure of example problem

Phase 1. Deciding Task Sequence

Phase 1.1 Order encoding for task sequence: A GA's structure and parameter settings affect its performance. However, the primary determinants of a GA's success or failure are the coding by which its genotypes represent candidate solutions and the interaction of the coding with the GA's recombination and mutation operators.

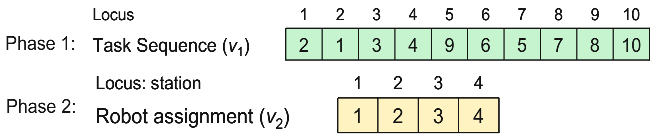


Fig. 4. Solution representation of a sample problem

A solution of the rALB problem can be represented by two integer vectors. Task sequence vector (v_1), which contains a permutation of assembly tasks, ordered according to their technological precedence sequence, and robot assignment vector (v_2). The solution representation method can be visually illustrated as in Figs. 3 and 4.

Phase 1.2 Reordering procedure: Considering precedence constraints, a task sequence may be infeasible since a precedent may appear before its successors. The reordering procedure will repair it into a feasible one before it is divided m parts to form a solution in the phase 1.

Phase 1.3 Breakpoint decoding: So, we use just one decoding procedure for phase 1, and the set of decoding procedure is to generate a feasible solution based on the task sequence and robot assignment schemes which are contained in the chromosomes. Then, the breakpoint decoding procedure inserts m points along the reordered task sequence vector to divide it into m parts, each of which corresponds to a station. The breakpoint decoding procedure consists of four main steps:

Step 1. Calculate the lower bound of the cycle time (C_{LB}) for the solution represented by the task sequence vector and robot assignment vector;

$$c_{LB} = \frac{1}{m} \sum_{i=1}^{n} \min_{1 \le l \le m} \{t_{il}\}; \tag{8}$$

Step 2. Find out a feasible cycle time as the upper bound of the cycle time (C_{UB});

Step 3. Find out the optimal cycle time via bisection method;

Step 4. Partition the task sequence into m parts with the optimal cycle time based on the robot assignment vector.

Here, a cycle time is said to be feasible if all the tasks can be allocated to the stations by allowing as many tasks as possible for each station under the constraint of the cycle time. The procedure to calculate the upper bound of cycle time is illustrated in procedure 2, and the bisection method to find out the optimal cycle time is shown in procedure 3. After calculating the optimal cycle time, it is easy to generate the breakpoints on the task sequence to divide it into m parts, each of which will correspond to a station, based on the robot assignment vector. Therefore, the tasks assigned for each station are determined according to a chromosome, and the cycle time can be calculated. An example of breakpoint decoding is showed in Fig. 5.

Phase 2. Assigning Robots to Station

Phase 2.1 Order encoding for robot assignment: A solution of the rALB problem can be represented by two integer vectors. In this phase, we code the robot assignment vector (v_2), which indicates the robot assigned for each station. The solution representation method can be visually illustrated as in Fig. 6. The coding space takes an exponential growth with the length of the chromosome; therefore, even one allele saved in the chromosome can decrease the coding space significantly. An obvious advantage to intentionally omit breakpoint vector, which is used in recently research, in the chromosome is that the coding space is dramatically decreased. As a result, the speed to find the global optimal is accelerated.

Phase 2.2 Breakpoint decoding: In this phase, this robot assignment vector (v_2) will decode into a solution with the task sequence vector (v_1), simultaneously by the breakpoint decoding procedure. It means when the tasks are assigned into each station, the robots will be assigned into each station at the same time. The illustration of the solution is shown in Phase 1.3.

Phase 3. Making a Schedule

Phase 3.1 Creating a schedule for the assembly line: We can make a schedule based on the chromosome (in Fig. 6) as fellow:

$$
\begin{aligned}
S = \{&(t_1, R_1 : 0-17), (t_2, R_1 : 17-38), (t_3, R_1 : 38-50), (t_4, R_2 : 50-71), \\
&(t_5, R_3 : 102-128), (t_6, R_2 : 71-89), (t_7, R_3 : 128-151), \\
&(t_8, R_4 : 151-176), (t_9, R_2 : 89-102), (t_{10}, R_4 : 176-202)\}
\end{aligned}
$$

Phase 3.2 Drawing a Gantt chart: Firstly, we draw a balancing chart for analyzing the solution (in Fig. 8). We can see the solution got the balancing for the line by comparing with the feasible solution from Fig. 2.

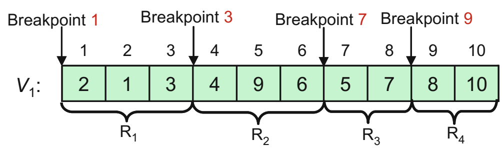


Fig. 5. Breakpoint decoding

In the real world, the assembly line is not just for producing one unit of the product. It should produce several units. So, we give the Gantt chart with 3 units for analyzing the solution just like Fig. 7, which is shown at the beginning. We can see the solution reduce the waiting time for the line by comparing with the feasible solution from Fig. 6. It also means the better solution got the balancing for the assembly line.

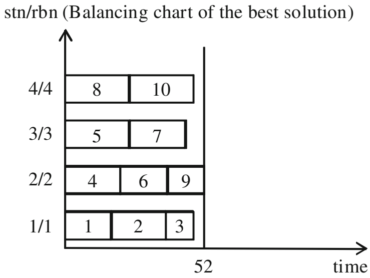


Fig. 6. The balancing chart of the best solution

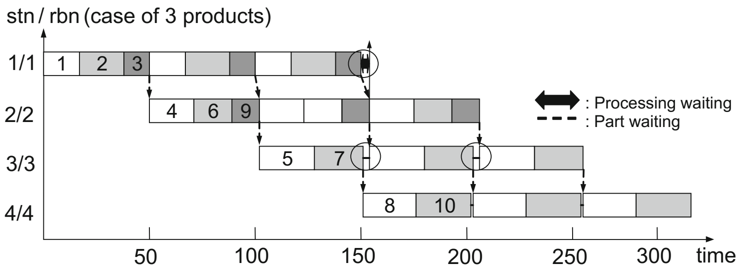


Fig. 7. Gantt chart for producing 3 units

3.1 Crossover Operators

In genetic algorithms, crossover recombines two chromosomes to generate a number of children. Offspring of crossover should represent solutions that combine substructures of their parental solutions [8]. In this paper, we use to kind of crossover methods: exchange crossover and mixed crossover. The detailed explanation of OX and PMX is introduced in [8].

(1) Exchange Crossover: A good schedule could be expected by exchanging task sequence and robot assignment schemes between a pair of parents. This kind of crossover is accomplished by selecting two parents and exchanging the task sequence vectors of the two parents to generate offspring.
(2) Mixed Crossover: The mixed order crossover consists of two crossover methods: Order Crossover (OX) and Partial-Mapped Crossover (PMX). For the rALB chromosome, the task sequence vector and the robot assignment vector are both of the style of permutation representation. Yet, the useful features of the task sequence vector and the robot assignment vector are different.

- For the task sequence vector, the useful feature of permutation representation to represent a rALB solution is the order information among the tasks.
- For the robot assignment vector, the acting feature is the number at each allele which indicates the robot no. assigned for a specific station.

3.2 Mutation Operators

In this study, two kinds of mutation operators are implemented: allele-based mutation and immigration mutation. For both the robot task sequence and robot assignment vectors, allele-based mutation randomly decides whether an allele should be mutated in a certain probability. Then, another position will be generated at random to perform exchange values with the mutated allele. In contrary to the canonical gene-by-gene mutation with very small probability at each generation, immigration mutation randomly generates one or more new members of the population from the same distribution as the initial population. This process prevents premature convergence of the population, and leads to a simple statement of convergence.

3.3 Fitness Evaluation

How to evaluate the fitness of chromosome is also key issue for GAs. In the most researches, objective function is used for fitness evaluation. However, the effective fitness function has to be different function and suitable to different problems. A central problem of fitness definition is how to guide to the most promising areas from the individuals of current generation. The characteristic of ALB problems are: ① We can calculate several different schedules with same cycle time. That means other conditions, such as idle time and total processing time are different, though the cycle time is same. For improvement of ALB, we also have to improve other conditions, not only cycle time. ② General, the solutions with better fitness of idle time and total processing time also can create better fitness of cycle time. We define the fitness function with considering not only the objective function (cycle time), but also idle time and total processing time. It is useful to guide to the most promising areas depending on the wide conditions from the individuals of current generation. We first define the fitness value $f_1(v)$ as the reciprocal of the cycle time $C_T(v)$ of the solution candidate v, the fitness values $f_2(v)$ as number of critical stations n_c and the fitness value $f_3(v)$ as the reciprocal of the standard deviation by the idle time of each station for balancing the workload of each station. t_k^W is the total processing time of each station. Then we give the adaptive weights for each fitness, where $w_1 + w_2 + w_3 = 1$. The special definition of fitness function is showed as follows:

$$f(v) = w(f(v)z^{\min}) + w(f(v)z^{\min}) + w(f(v)z^{\min}), \tag{9}$$

where

$$f_1(v) = \frac{1}{C_T(v)}, \tag{10}$$

$$f_2(v) = n_c(v), \tag{11}$$

$$f_3(v) = \frac{1}{m-1}\sum_{k=1}^{m}((C_T(v) - t_k^W(v)) - \frac{1}{m}\sum_{j=1}^{m}(C_T(v) - t_j^W(v))). \tag{12}$$

4 Local Search Procedure

Two individuals with high fitness values are likely to have dissimilar machine assignment and operation sequences, and the recombination may result in offspring of poor performance. This means that the genetic operations by themselves have limited ability in finding the global optimal solution. Many researchers have found that the convergence speed of simple GAs is relatively slow [7]. One promising approach for improving the convergence speed to the global optimum is the use of local search in GAs [17]. In this study, two local search methods are proposed to enhance the search ability of GA. Both the two local search methods are based on critical stations in order to improve their effectiveness and efficiency.

4.1 Task Sequence Search

Let $s(i)$ be the station to which task i is assigned. First we introduce the logical function $W_i(k)$ which return false if task i cannot be transferred from station $s(i)$ to station k, and true otherwise:

$$\begin{aligned} W_i(k) = &\text{false if} \\ &\quad s(i) > k \text{ and } j \in \mathrm{suc}(i), s(j) > k, \text{ or} \\ &\quad s(i) < k \text{ and } j; i \in \mathrm{suc}(j), s(j) < k \\ W_i(k) = &\text{true, otherwise.} \end{aligned}$$

For a pair of tasks i, j, if

$$W_i(s(j)) = W_j(s(i)) = \text{true} \wedge i \notin \mathrm{pre}(i)\ j \notin \mathrm{pre}(i),$$

then, the two tasks are said to be exchangeable. Let $R(i)$ be the robot which is allocated for station $s(i)$, and Tl be the total assembly time of station l. For a critical task i and task j which is not in $s(i)$, if

$$t_{j,s(i)} < t_{i,s(i)} \text{ and } T_s(j) - t_{j,s(j)} + t_{i,s(j)} < C_T,$$

then the exchange between tasks i and j is called worthwhile. The task sequence search is as follows:

Step 1. Decode the chromosome and identify a critical station l^*;
Step 2. For a task i that assigned to the critical station l^*, look for a task j that is not assigned to station l^*, and is exchangeable with task i.
Step 3. Repeat Step 2 until the exchange between tasks i and j is worthwhile or all the tasks in the critical station are tried.
Step 4. Repeatedly perform the above search steps until the chromosome cannot be improved any more.

4.2 Robot Assignment Search

Let $N(i)$ denote the set of machine assignment neighborhood of solution i. The enlarged two-pace machine assignment neighborhood is defined as the union of the neighborhood of each robot assignment neighbor of solution i

$$N^2(i) = \cup_{j \in N(i)} N(j). \tag{13}$$

During the robot assignment search, the local search will implement over two-pace neighborhood when it reaches the local optima of one-pace neighborhood, and is called two-pace robot assignment search. During the robot assignment search, when the local optima of two-pace robot assignment neighborhood is reached, the neighbors of the two-pace local optima are improved by the task sequence search to help the robot assignment search escape from the local optima. Since the size of the two-pace neighborhood is very large and the task sequence search is very computationally complex, to improve each neighbor solution by task sequence search needs a great computation time. Therefore, only those neighbor solutions that are local optimums of the two-pace robot assignment neighborhood are improved by task sequence search so as to save computation time.

5 Numerical Experiments

In the literature, no benchmark data sets are available for rALB. We collect 8 representative precedence graphs from [1], which are widely used in the sALB-1 literature [25]. These precedence graphs contain with 25-297 tasks. From each precedence graph, 4 different rALB-2 problems are generated by using different WEST ratios: 3, 5, 7, 10, 15. WEST ratio, as defined by Dar-EI [19], measures the average number of activities per station. For each problem, the number of station is equal to the number of robots, and each task can be processed on any robot. The task time data are generated at random, while two statistical dependence are maintained: (1) statistical dependence of task times on the task type, (2) statistical dependence of task times on the robot on which the task is processed.

To validate the effectiveness of our hybrid GA, we compare our approach with Levitin et al.'s two algorithms. Recently, Levitin et al. [18] develop an efficient approach for rALB problems. It is similar research with our purpose, aims to achieve a balanced distribution of work between different stations (balance the line) while assigning to each station the robot best fit for the activities assigned to it. Levitin et al. proposed two algorithms named as recursive assignment method and consecutive assignment method. If we look each robot as one type and deleting this type from the total types of robots when the robot is allocated for a station, Levitin et al.'s two algorithms can be adapted to solve the 32 rALB-2 problems here. Hence, the two algorithms proposed by Levitin et al. are also used to solve the 32 problems. The adopted parameters of the hGA are listed as following: maximal generations maxGen = 1000; population size popSize = 100;

Table 1. Performance of the proposed algorithm

Test problems			Cycle time (C_T)			Test problems			Cycle time (C_T)		
♯ of task	♯ of station	West ratio	Levitin et al.'s recursive	Levitin et al.'s consecutive	Proposed approach	♯ of task	♯ of station	West ratio	Levitin et al.'s recursive	Levitin et al.'s consecutive	Proposed approach
25	3	8.33	518	503	503	89	8	11.13	638	505	494
	4	6.25	351	330	327		12	7.42	455	371	370
	6	4.17	343	234	213		16	5.56	292	246	236
	9	2.78	138	125	123		21	4.24	277	209	205
35	4	8.75	551	450	449	111	9	12.33	695	586	557
	5	7.00	385	352	344		13	8.54	401	339	319
	7	5.00	250	222	222		17	6.53	322	257	257
	12	2.92	178	120	113		22	5.05	265	209	192
53	5	10.60	903	565	554	148	10	14.80	708	638	600
	7	7.57	390	342	320		14	10.57	537	441	427
	10	5.30	35	251	230		21	7.05	404	325	300
	14	3.79	243	166	162		29	5.10	249	210	202
70	7	10.00	546	490	449	297	19	15.63	1129	674	646
	10	7.00	313	287	272		29	10.24	571	444	430
	14	5.00	231	213	204		38	7.82	442	348	344
	19	3.68	198	167	154		50	5.94	363	275	256

exchange crossover probability $p_C = 0.20$; mixed crossover probability $p_M^1 = 0.80$; task sequence mutation probability $p_M^2 = 0.05$; robot assign. mutation probability $p_M^3 = 0.10$. Levitin et al.'s two algorithms used the same population size and maximal generations as the hGA.

Table 1 shows the experiment results of Levitin et al.'s two algorithms and proposed approach for 32 different scale test problems. As depicted in Table 1, all of results by proposed approach are obvious better than Levitin et al.'s recursive assignment method. And most results of proposed approach are better than and Levitin et al.'s consecutive assignment method, except to the test 25-3, test 35-7 and test 111-17. The differences of 32 test problems between proposed approach with Levitin et al.'s approaches are showed in Fig. 8.

To clearly see the convergence performance of proposed approach, we give the evolutionary process of proposed approach and Levitin et al.'s two methods by combining 3 different scale test problems: small-scale test 35-12, middle-scale test 111-17 and large-scale test 148-21. The evolutionary processes are showed in Fig. 9. As depicted in Fig. 9, the convergence of proposed approach is faster than each of other approaches. The difference of convergence increased with increasing the test problems size. Furthermore, though Levitin et al.'s consecutive found the same result with proposed approach for test 111-17, but the convergence speed of proposed approach was faster. The simulations show that propose hybrid GA with combining local search approach is computationally efficient and effective to find the best solution and convergence speed.

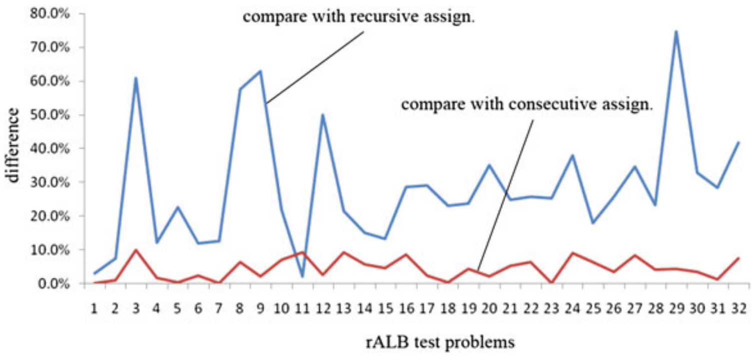
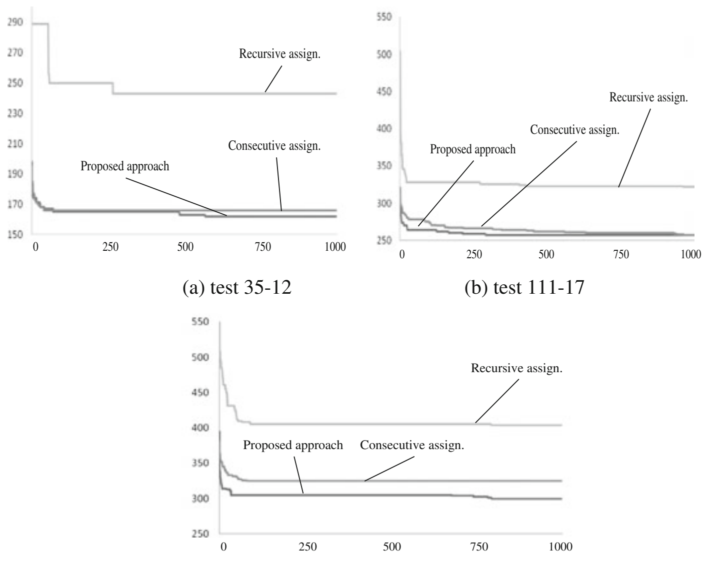


Fig. 8. The difference between different approaches



(a) test 35-12 (b) test 111-17

(c) test 148-21

Fig. 9. Evolutionary process of 3 different approaches

6 Conclusion

Robotic assembly line has been an important manufacturing system in the modern era. The objective of this work is to develop an efficient solution for the robotic assembly line balancing problem. This solution aims to achieve a balanced distribution of work between different stations and assign to each station the robot best fit for the activities assigned to it. The result of such solution would be an increased production rate of the line.

A new representation method adapting the GA to the rALB-II problem is proposed. Advanced genetic operators adapted to the specific chromosome structure and the characteristics of the rALB problem are used. In order to strengthen the search ability, two kinds of local search are integrated under the framework the genetic algorithm. The coordination among the three kinds of local search is well considered. The neighborhood structure of the local search can be adjusted dynamically. The balance between genetic search and local search is also investigated. The performance of proposed method is validated through simulation experiments. The simulation shows that our algorithm is computationally efficient and effective to find the best solution. The solution obtained by our algorithm outperforms the results from previous works.

Acknowledgements. This work is partly supported by the National Natural Science Foundation of China under Grant 61572100, and in part by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS) No. 15K00357.

References

1. ALB: Assembly line balancing: Data sets & research topics (2016). http://www.assembly-line-balancing.de/
2. Anderson EJ, Ferris MC (1994) Genetic algorithms for combinatorial optimization: the assemble line balancing problem. Informs J Comput 6(2):161–173
3. Bautista J, Suarez R et al (2000) Local search heuristics for the assembly line balancing problem with incompatibilities between tasks. In: IEEE International Conference on Robotics and Automation, 2000, Proceedings, ICRA, pp 2404–2409
4. Brown EC, Sumichrast RT (2005) Evaluating performance advantages of grouping genetic algorithms. Eng Appl Artif Intell 18(1):1–12
5. Daoud S, Chehade H et al (2014) Solving a robotic assembly line balancing problem using efficient hybrid methods. J Heuristics 20(3):235–259
6. Falkenauer E, Delchambre A (1997) A genetic algorithm for bin packing and line balancing. In: IEEE International Conference on Robotics and Automation, 1992, Proceedings, vol 2, pp 1186–1192
7. Gao J, Sun L et al (2009) An efficient approach for type II robotic assembly line balancing problems. Comput Ind Eng 56(3):1065–1080
8. Gen M, Cheng R (1997) Genetic Algorithms and Engineering Design. Wiley, New York
9. Gen M, Cheng R (2000) Genetic Algorithms and Engineering Optimization. Wiley, New York
10. Gen M, Tsujimura Y, Li Y (1996) Fuzzy assembly line balancing using genetic algorithms. Comput Ind Eng 31(31):631–634
11. Gonçalves JF, Almeida JRD (2002) A hybrid genetic algorithm for assembly line balancing. J Heuristics 8(6):629–642
12. Gutjahr AL, Nemhauser GL (1964) An algorithm for the line balancing problem. Manage Sci 11(2):308–315
13. Helgeson WB, Salveson ME, Smith WW (1954) How to Balance an Assembly Line. Technical Report Carr Press, New Caraan
14. Janardhanan MN, Nielsen P, Ponnambalam SG (2016) Application of particle swarm optimization to maximize efficiency of straight and U-shaped robotic assembly lines. Springer
15. Kim YK, Kim YJ, Kim Y (1996) Genetic algorithms for assembly line balancing with various objectives. Comput Ind Eng 30(3):397–409

16. Kock S, Vittor T et al (2011) Robot concept for scalable, flexible assembly automation: a technology study on a harmless dual-armed robot. In: IEEE International Symposium on Assembly and Manufacturing, pp 1–5
17. Krasnogor N, Smith J (2000) A memetic algorithm with self-adaptive local search: TSP as a case study. In: Proceedings of 2000 Genetic and Evolutionary Computation Conference, pp 987–994
18. Levitin G, Rubinovitz J, Shnits B (2006) A genetic algorithm for robotic assembly line balancing. Eur J Oper Res 168(3):811–825
19. (Mansoor) EMDE (1973) Malbła heuristic technique for balancing large single-model assembly lines. IIE Trans 5(4):343–356
20. Mukund Nilakantan J, Ponnambalam SG et al (2015) Bio-inspired search algorithms to solve robotic assembly line balancing problems. Neural Comput Appl 26(6):1379–1393
21. Nearchou AC (2007) Balancing large assembly lines by a new heuristic based on differential evolution method. Int J Adv Manufact Technol 34(9):1016–1029
22. Rubinovitz J, Levitin G (1995) Genetic algorithm for assembly line balancing. Int J Prod Econ 41(41):343–354
23. Rubinovitz J, Levitin G (1995) Genetic algorithm for line balancing. Int J Prod Econ 41:343–354
24. Sabuncuoglu I, Erel E, Tanyer M (2000) Assembly line balancing using genetic algorithms. J Intell Manufact 11(3):295–310
25. Scholl A (1993) Data of assembly line balancing problems. Schriften zur Quantitativen Betriebswirtschaftslehre
26. Scholl A (1999) Balancing and Sequencing of Assembly Lines. Physica-Verlag, Heidelberg
27. Scholl A, Fliedner M, Boysen N (2010) Absalom: balancing assembly lines with assignment restrictions. Eur J Oper Res 200(3):688–701
28. Tsujimura Y, Gen M, Kubota E (1995) Solving fuzzy assembly-line balancing problem with genetic algorithms. In: International Conference on Computers and Industrial Engineering, pp 543–547
29. Yoosefelahi A, Aminnayeri M et al (2012) Type II robotic assembly line balancing problem: an evolution strategies algorithm for a multi-objective model. J Manufact Syst 31(2):139–151

A Hybrid Model of AdaBoost and Back-Propagation Neural Network for Credit Scoring

Feng Shen[1(✉)], Xingchao Zhao[1], Dao Lan[2], and Limei Ou[2]

[1] School of Finance, Southwestern University of Finance and Economics, Chengdu 611130, People's Republic of China
shenfeng1213@gmail.com
[2] Southwestern University of Finance and Economics, Chengdu 611130, People's Republic of China

Abstract. Owing to the development of internet finance in China, credit scoring is growing into one of the most important issues in the field of financial risk management. Quantitative credit scoring models are widely used tools for credit risk assessment in financial institutions. In this study, an AdaBoost algorithm model based on back-propagation neural network for credit scoring with high accuracy and efficiency is proposed. We first illustrate the basic concepts of back-propagation neural network and AdaBoost algorithm and propose a hybrid model of AdaBoost and back-propagation neural network, then two real-world credit data sets are selected to demonstrate the effectiveness and feasibility of the proposed model. The results show that the proposed model can get higher accuracy compared to other classifiers listed in this study.

Keywords: Credit scoring · AdaBoost model · Back-propagation neural network

1 Introduction

Credit scoring has grown an increasingly important issue of financial risk management in financial institutions since the 2008 financial crisis. It often calculates by following a set of decision models and other underlying technologies, does a great favor for lenders' judging whether an application of credit should be approved or rejected [27]. When some applicants fail to repay their debt, it leads to a direct economic loss for the lenders. In addition, the sub-prime mortgage crisis occurred in the USA has caused some financial institutions loss billions of dollars due to customers' default. However, if a credit-granting-institution rejects all applicants of loans even with good credit scores, it will suffer the potential revenues it can earn from the applicants in the future. Therefore, an efficient decision support with high accuracy becomes a clear need for financial institution.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_6

Due to the great importance of credit risk assessment, an increasing research stream is focusing upon credit risk assessment and credit scoring using different methods and models, approaches such as linear discriminant analysis (LDA) [1], logistic regression analysis LRA [26], k-nearest neighbor (KNN) [12,13] and decision tree [5] which uses the statistics disciplines. With the development of artificial intelligence(AI) techniques, artificial neural network (ANN) [3,8,14,21, 30,31], support vector machines (SVM) [2,11,25,34], genetic algorithm (GA) [7,22,28], rough set [4] and case-based reasoning [15–19] are used for credit scoring. Also some combined and ensemble approaches perform quite well with high accuracy and efficiency, including fuzzy system and artificial neural network [20,23], rough set and support vector machine [33], fuzzy system and support vector machines [29], case-based reasoning and support vector machines [19], neural network ensemble [32] etc.

Inspired by the combined and ensemble theories, this study attempts to purpose an ensemble AdaBoost model based on BP neural network for credit scoring. The main idea of AdaBoost algorithm is to maintain a distribution of weights over the training samples and adjust them after each basic classifier sorting cycle adaptively. In our study, we use ten single back propagation neural networks as the weak learners contends of a three-layer feed-forward network each, and the final strong ensemble classifier is constructed by the method of weighted voting associated with the performances of the weak learners. Based on the experimental results achieved from two public available credit data sets, our proposed AdaBoost model based on BP neural network obtains a good performance with higher accuracy efficiency compared to other models used in this study, which indicates a wide prospect usage in financial risk management.

The rest of this paper is organized as follows. Section 2 gives a brief formulation concerns about back propagation neural network and AdaBoost algorithm. Section 3 presents a hybrid model of AdaBoost and BP neural network. To verify the accuracy and effectiveness of the purposed model, empirical validation results of the model using the German credit data set and Australian credit data set are analyzed in Sect. 4. Finally, a short conclusion and discussion are presented in Sect. 5.

2 Methodology Formulation

2.1 Back-Propagation Neural Network Theory

Back-propagation neural network model was proposed by Rumelhart and Mccelland in 1985 [24], it is a kind of multi-layered forward feed type error counterbiography neural network. This model usually consists of input layer, output layer and hidden layer. Each layer includes certain nodes, and each node expresses a neuron. There exists interconnection among the nodes of different layers, but there is no connection among nodes in the same layer. Among them, the single hidden layer BP network's application is the most common. The typical structure of BP neural network is as shown in Fig. 1.

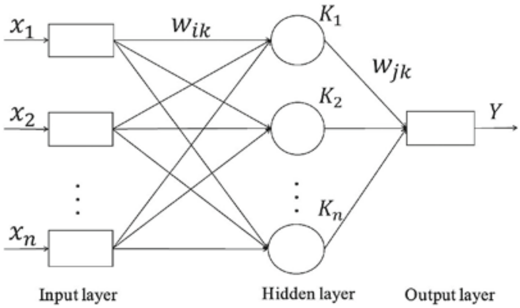


Fig. 1. Structure of single hidden layer BP neural network

The core thought of BP neural network can be described as following: training signal forward-propagating and error signal reverse dissemination. During the process of forward-propagating, the signal spreads from input layer, dealt with connection weights between input layer and hidden layer, then is transformed by the active function of hidden layer and transmitted to the output layer. If expected output can't be achieved from the output layer, then another process of error back propagation begins, which would continuously adjust the weights and bias of each layer in order to ultimately minimize the error of the network. This kind of signal forward-propagating and error back-propagating is to go iteration and iteration till the system output error is reduced to an acceptance degree or obtains the network constraints. Its concrete process is shown in Fig. 2.

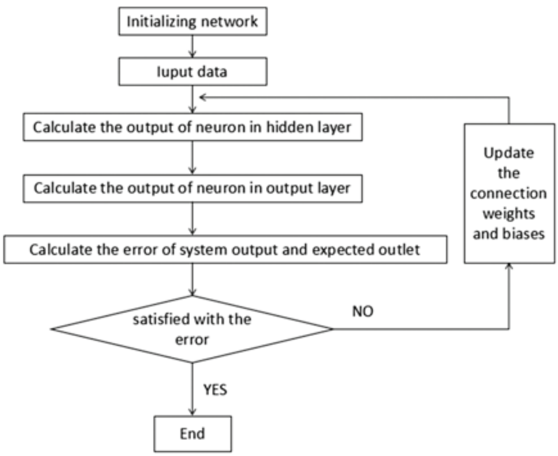


Fig. 2. Process of BP neural network model

2.2 Basic Theory of AdaBoost Algorithm

AdaBoost is a very popular boosting algorithm for binary classification, and it solved many practical difficulties of the earlier boosting algorithms. It was initially proposed by Yoav Freund and Robert Schapire in 1997 [10]. This algorithm constructs a high-quality compositive classifier by combining trained weak classifiers sequentially while putting more attention on those weak classifiers accompanied with good performance. There are several methods for combining results from base weak learners into one stronger predictor. Uniform voting, distribution summation, Bayesian combination, etc. are widely used for ensemble algorithms. Adaptive boosting algorithm has been applied in many fields such as speech recognition [9], moving vehicle classification based on images [6] and etc.

The main ideas of AdaBoost algorithm is to maintain a distribution of weights over the training samples and adjust them after each basic classifier sorting cycle adaptively. The weights of training data which are wrongly classified by current weak learner will be increased, otherwise, decreased if the samples are correctly classified. During the training process, the prediction error of the weak classifiers should be less than 0.5. And the voting weights of base classifiers will be increased while the decreasing of the prediction error, which means a larger voting weight this weak learner will take in the final ensemble output. Pseudocode for AdaBoost algorithm is shown in Fig. 3. As a series of base learners have been achieved, AdaBoost calculates a value α_t that is assigned to h_t, and the final hypothesis H is constructed via T weak classifiers using a weighted voting ensemble method.

AdaBoost Algorithm:

Given a training data set: $T = \{(x_1, y_1), \dots (x_i, y_i), \dots (x_n, y_n)\}$ $where$ $x_i \epsilon X, y_i \epsilon Y = \{-1, +1\}$

Initialize Distribution: $D_1(i) = \frac{1}{n}$, $i = 1,2, \dots, n$

For t = 1,2, … T:

•Train weak classifier using samples with distribution D_t

•Get weak hypothesis $h_t : X \to \{-1, +1\}$

•Calculate the error of h_t:

$$\varepsilon_t = \sum_{i=1}^{n} D_t(i) I_t(\mathrm{i})$$

$$\text{Where } I_t(\mathrm{i}) = \begin{cases} 1 & if \;\; h_t \neq y_i \\ 0 & if \;\; h_t = y_i \end{cases}$$

• Calculate the weight of current hypothesis: $\alpha_t = \frac{1}{2} * \ln(\frac{1-\varepsilon_t}{\varepsilon_t})$

•Update distribution: $D_{t+1}(i) = \frac{D_t(i)}{Z_t} \exp(-\alpha_t * y_i * h_t(xi))$

$$\text{Where } Z_t = \sum_{i=1}^{n} D_t(i) * \exp(-\alpha_t * y_i * h_t(xi))$$

Output final classifier:

$$H(x) = sign(\sum_{t=1}^{T} \alpha_t * h_t(x_i))$$

Fig. 3. Pseudocode for AdaBoost algorithm

3 Hybrid Model of AdaBoost and BP Neural Network

In this section, a hybrid model is constructed based on AdaBoost and BP neural network for credit scoring. According to the AdaBoost algorithm, the efficiency and accuracy of the ensemble predictor based on a series of weak classifiers largely depends on the accuracy of those base learners, the correlation coefficient of the ensemble predictor and the base classifiers is directly proportional. As BP neural network model is very fledged both in theory and practical applications with high accuracy and efficiency compared to other widely used models. So, in this study, BP neural network is selected as the weak classifier, and as the final ensemble classifier requires a certain method to combine those weak learners into an efficient ensemble system, we adopt the weighted voting method to combine the outputs of each weak classifier for the final outputs. The frame and pseudocode of hybrid model based on AdaBoost and BP neural network are shown in Figs. 4 and 5. The main steps used to construct ensemble predictor are shown as follows.

Step 1. Select the training samples randomly from the chosen databases with the method of 10-fold cross-validation, and each group training sample is assigned with the same weights:

$$\mathrm{T} = \{(x1, y1), \cdots, (xi, yi), \cdots, (xn, yn)\} \text{ while } xi \in X, yi \in Y = \{-1, +1\}.$$

The weights of this distribution on training samples on round t is denoted $D1(i)$. Initially, all weights are assigned with equal value $D1(i)$:

$$D1(i) = \frac{1}{n}, i = 1, 2, \cdots, n, \tag{1}$$

where n represents the total number of samples in the training data set;

Step 2. Parameters of the BP neural network, which include the structure of the network, neurons of each layer, expected goal, are set up according to the attributes of the input and output training samples. Then initialize the weights and biases of BP neural network.

Step 3. Training the BP neural network with the processed databases:

Step 3.1 Calculate input and output values of hidden layer: net_j, b_j

$$\mathrm{net}_j = \sum_{i=1}^{n} w_{ij} \times x_i - \theta_j, \tag{2}$$

$$b_j = f(\mathrm{net}_j), \tag{3}$$

where w_{ij} is the connection weights between input layer and hidden layer and θ_j represents the Threshold values of hidden layer, $f(x)$ represents the transfer function Sigmoid. Namely:

$$f(\mathrm{net}) = \frac{1}{1 + e^{-\mathrm{net}}}. \tag{4}$$

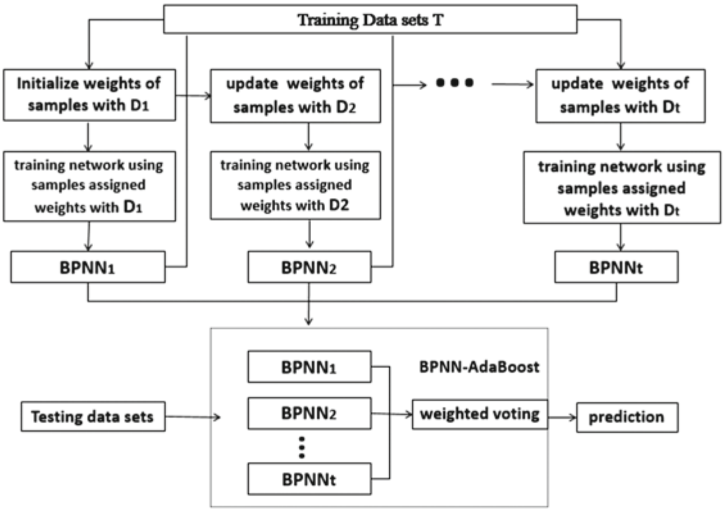


Fig. 4. The frame of hybrid model of AdaBoost and BP neural network

AdaBoost Model Based on BP neural network

Input: Training dataset $T = \{(x_1,y_1),\dots(x_i,y_i),\dots(x_n,y_n)\}$, where $x \in R^n, y_i \in \{-1,+1\}$
Weak classifier training algorithm (Back-Propagation algorithm)
Number of Learning Round: T
Initialize Distribution: $D_1(i) = \frac{1}{n}$, $i = 1,2,\dots,n$
For t=1: T Do

1. Select the subset training data drawn from the distribution $D_1(i)$.
2. Train the base classifier with the samples and obtain the hypothesis $h_t, h_t \in \{-1,+1\}$.
3. Calculate the classifying error of $h_t(x)$:

   $\varepsilon_t = \sum_{i=1}^{n} D_t(i) I(h_t(x_i))$ where $I(h_t) = \begin{cases} 1 & if\ h_t(x_i) \neq y_i \\ 0 & if\ h_t(x_i) = y_i \end{cases}$
4. Calculate the weight of current hypothesis h_t

   $\alpha_t = \frac{1}{2}\ln(\frac{1-\varepsilon_t}{\varepsilon_t})$
5. Update distribution $D_t(i)$

   $D_{t+1}(i) = \frac{D_t(i)}{Z_t}\exp(-\alpha_t * y_i * h_t(x_i))$

   Where Z_t is normalization constant for $D_{t+1}(i)$ become a proper distribution.

Output: Construct the final classifier based on the weak classifiers that have been trained using weighted voting method.

$$H(x) = sign(\sum_{t=1}^{T} \alpha_t * h_t(x))$$

Fig. 5. Pseudocode of hybrid model of AdaBoost and BP neural network

Step 3.2 Calculate input and output values of output layer:

$$\text{net}_i = \sum_{j=1}^{m} v_{jt} \times b_j - \gamma_i, \tag{5}$$

$$o_i = f(\text{net}_i), \tag{6}$$

where v_{jt} is the connection weights between hidden layer and output layer and γ_i represents the Threshold values of output layer. In this study, the hidden layer contains m neurons.

Step 3.3 Calculate the error d_i of each neuron of output using network's goal vector and its response output:

$$d_i = y_i - o_i, \quad (i = 1, 2, \cdots, n). \tag{7}$$

Step 3.4 Calculate the total error of the network $e(k)$:

$$e(k) = \frac{1}{2}\sum_{i=1}^{n}(d_i)^2. \tag{8}$$

Step 3.5 If the total error above is acceptable, the process stops. Otherwise, revise and. There are many ways of weight changes, we use the most common gradient descent method:

$$w_{ij}^{k+1} = w_{ij}^{k} + \Delta w_{ij}^{k} + \partial \Delta w_{ij}^{k-1}, \tag{9}$$

$$w_{jt}^{k+1} = w_{jt}^{k} + \Delta w_{jt}^{k} + \partial \Delta w_{jt}^{k-1}. \tag{10}$$

This process continuous till the system output error is reduced to an acceptance degree or obtains the network constraints.

Step 4. When training epochs go to round t, a base classifier $h_t(x)$ could be achieved with the weights distribution D_t and then calculate the weighted error ε_t from model:

$$\varepsilon_t = \sum_{i=1}^{n} D_t(i) \times I(h_t(x_i)), \tag{11}$$

where I$(h_t(x_i))$ is the indicator function, and its mathematical expression shows as below:

$$I(ht) = \begin{cases} 1, & \text{if } h_t(x_i) \neq y_i, \\ 0, & \text{if } h_t(x_i) = y_i. \end{cases} \tag{12}$$

Step 5. Calculate the weight of base classifier $h_t(x)$

$$a_t = \frac{1}{2}\ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right). \tag{13}$$

As is shown above, if the prediction error of the samples is less than 0.5, the weight is bigger than zero, and with the decreasing of the prediction error, will be increased meanwhile, which means a larger voting weight of this weak learner in the final ensemble output.

Step 6. Update the weights of each samples in the training databases for the next iteration:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t}\exp(-a_t \times y_i \times h_t(x_i)), \tag{14}$$

where Z_t is normalization constant for $D_{t+1}(i)$ become a proper distribution:

$$Z_t = \sum_{i=1}^{n} D_t(i) \times \exp(-a_t \times y_i \times h_t(x_i)). \tag{15}$$

In this way, the weights of training data which are wrongly classified by current weak learner will be increased or decreased if the samples are correctly classified.

Step 7. As the training iteration goes on, a series of base classifiers will be obtained, then combine these classifiers during the processes:

$$f(x) = \sum_{t=1}^{T} a_t \times h_t(x). \tag{16}$$

A series of base classifiers with diversity are collected with loops from step 2 to step 5. Finally, an AdaBoost model based on the weak learners trained above is constructed:

$$H(x) = \mathrm{sign}(\sum_{t=1}^{T} a_t \times h_t(x)) \tag{17}$$

4 Empirical Analysis

In order to test the performance of the proposed AdaBoost algorithm based on BP neural network model, two available real world credit data sets with detailed input attributes description (German and Australian credit data sets) are used in this study. These two credit databases are available from open access UCI Machine Learning Repository. The German credit data set consists of 1000 instances with 700 samples labeled as creditworthy and 300 samples classified as poor credit. For each instance, there are 24 input variables described 19 attributes with 4 attributes changed to dummy variables. Also, the 25th variable is the label of the instance with two numerical descriptions, and label 1 stands for a worthy credit assessment of current instance, label 2, on the contrary, represents a bad credit evaluation. The meaning of original attributes are described in Table 1; The Australian credit data set is interesting because there is a good mix of attributes – continuous, nominal with small numbers of values, and nominal with larger numbers of values, which includes 690 instances in total with 383 samples classified as trustworthy and 307 samples labeled as poor credit. For each instance, 15 variables described 15 features of personal information and financial history of applicants, the last feature is labeled as approved (marked as 0) or rejected (marked as 1). All attribute names and values have been changed to meaningless symbols to protect confidentiality of the data.

Table 1. Original attributes in the German credit data set

Number	Description	Class
Attribute 1	Status of existing checking account	Qualitative
Attribute 2	Duration in month	Numerical
Attribute 3	Credit history	Qualitative
Attribute 4	Purpose	Qualitative
Attribute 5	Credit amount	Numerical
Attribute 6	Savings account/bonds	Qualitative
Attribute 7	Present employment since	Qualitative
Attribute 8	Instalment rate in percentage of disposable income	Numerical
Attribute 9	Personal status and sex	Qualitative
Attribute 10	Other guarantors	Qualitative
Attribute 11	Present residence since	Numerical
Attribute 12	Property	Qualitative
Attribute 13	Age in years	Numerical
Attribute 14	Other instalment plans	Qualitative
Attribute 15	Housing	Qualitative
Attribute 16	Number of existing credits at this bank	Numerical
Attribute 17	Job	Qualitative
Attribute 18	Number of people being liable to provide maintenance for	Numerical
Attribute 19	Telephone	Qualitative
Attribute 20	Foreign worker	Qualitative

4.1 Experiment Design

In the experiment, the AdaBoost algorithm model based on BP neural network is used. The final classifier contents ten weak learners constructed with a three-layer feed-forward BP neural network model. As to the neurons in hidden of each weak learner, 9 nodes in hidden layer dealt with German data set and 20 nodes in hidden layer in Australian data set are set. For comparison purpose, some commonly used models, such as linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), logistic analysis (Log A), K-Nearest neighbor (KNN), artificial neural network (ANN), back propagation neural network (BPNN), and least squares support vector machine (LSSVM) are also applied. In this study, training samples are selected by the method of 10-fold cross-validation to determine the model parameters. In BP neural network model, a common three-layer feed forward net trained by Levenberg-Marquartdt algorithm is employed, the transfer function in hidden layer is applied with Sigmoid function, and with Purelin function in output layer. In addition, three evaluation criteria, Type I Accuracy, Type II Accuracy and Total Accuracy are used, which are defined as follows [34]:

$$\text{Type I Accuracy} = \text{Specificity} = \frac{\text{number of both observed bad and classified as bad}}{\text{number of observed bad}}, \quad (18)$$

$$\text{Type II Accuracy} = \text{Sensitivity} = \frac{\text{number of both observed good and classified as good}}{\text{number of observed good}}, \quad (19)$$

$$\text{Total Accuracy} = \frac{\text{number of correct classification}}{\text{number of evaluation sample}}. \quad (20)$$

4.2 Experiment Result

As is shown in Tables 2 and 3, validation results for the German credit data set and Australian data set are achieved using the different algorithms and settings described above. The classification results of the first six classifiers in Tables 2 and 3 are from another article [35]. From Tables 2 and 3, several interesting findings can be drawn.

First of all, the total accuracy of the proposed AdaBoost model based on BP neural network is much better than other classifiers in both German and Australian data sets. Followed by the least squares support vector machine. Probable explaining reasons of the high accuracy achieved by AdaBoost model includes two aspects. For one thing, the accuracy of AdaBoost model depends much on the accuracy of the weak learners which it ensembles, and as we can see from Tables 2 and 3, BP neural network do a good credit classification compared to other commonly used models. For another reason, the main idea of AdaBoost system gives larger weights to those base classifiers with small prediction error, and as a result, the total accuracy of AdaBoost model get the best performance among all the classifiers compared in this study.

Second, referring to Type I Accuracy, the presented AdaBoost model based on BP neural network performs the best, compared to other classifiers employed in this study both in German credit data set and Australian credit data set. Followed by the linear discriminant analysis model in German data set and logistic analysis model in Australian data set.

Third, in terms of Type I Accuracy and Type II Accuracy, all models used in this study perform better with higher accuracy in Type II Accuracy than Type I Accuracy in general, which indicates that it's more difficult to classify customers with bad credit conditions from those of good credit evaluation cause of the complexity of credit risk.

Fourth, according to the Type II Accuracy shown in Table 2, our proposed model only ranks the fourth compared to other models for comparison. Supposed reason of this phenomenon may be the default of the AdaBoost algorithm, there is no exotic existence of noisy data in German data set, which will attract more attention of our AdaBoost model based on BP neural network on these noisy signals, as a result, the accuracy of our proposed model may not be satisfied as we wish and thus performs ordinarily in German credit data set.

Finally, compared with Tables 2 and 3, it's easy to find out that the performance of the German data set is worse than that of the Australian data set.

Table 2. Credit evaluation results comparison of different models for German credit dataset

Models	Type I Accuracy		Type II Accuracy		Total Accuracy	
	%	Rank	%	Rank	%	Rank
LDA	72	2	74.57	6	73.8	5
QDA	66.57	3	69.33	8	67.4	8
Log A	50.33	5	88.14	3	76.8	3
KNN	27	8	90.57	1	71.5	6
ANN	46.89	7	73.46	7	69.43	7
BPNN	63.83	4	79.36	5	75.8	4
LSSVM	49.67	6	88.86	2	77.1	2
BPNN_AdaBoost	78.7	1	84.19	4	83	1

Table 3. Credit evaluation results comparison of different models for Australian credit dataset

Models	Type I Accuracy		Type II Accuracy		Total Accuracy	
	%	Rank	%	Rank	%	Rank
LDA	80.94	6	92.18	2	85.94	5
QDA	66.12	8	91.38	3	80.14	6
Log A	85.9	2	86.32	6	86.09	4
KNN	81.72	5	54.4	8	69.57	8
ANN	72.56	7	83.61	7	78.94	7
BPNN	83.95	4	87.65	5	86.23	3
LSSVM	85.12	3	89.25	4	86.96	2
BPNN_AdaBoost	89.06	1	94.59	1	92.03	1

There are two possible reasons. On one hand, the credit market in Germany is more complex than that in Australia. On the other hand, there is more non-linearity in the German data set than in the Australian data set. Overviewing the above performance, one interesting phenomenon shows that there is a great capability raising of all three evaluation criteria both in the German data set and Australian data set compared BP neural network with the AdaBoost model based on BP neural network, from the viewpoint of Total Accuracy, there is a 9.5% raising in the German data set and a 6.7% raising in the Australian data set.

5 Conclusion

In this paper, a hybrid model of AdaBoost and BP neural network is proposed for credit risk evaluation. According to the empirical results, we find that our proposed hybrid model is the best one compared with other seven models for two publicly available credit data sets, which indicates that our proposed hybrid model of AdaBoost and BP neural network has a good practicability for credit scoring. In the future, we will use other method as base classifiers, for example, support vector machine and decision tree, and research other ensemble algorithm like bagging algorithm for credit risk assessment.

Acknowledgements. This research was supported by the project of Research Center for System Sciences and Enterprise Development (Grant No.Xq16C03).

References

1. Altman EI (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J Finan 23(4):589–609
2. Baesens B, van Gestel T et al (2003) Benchmarking state-of-art classification algorithms for credit scoring. Oper Res Soc 54(6):627–635
3. Baesens B, Setiono R et al (2003) Using neural network rule extraction and decision tables for credit-risk evaluation. Manag Sci 49(3):312–329
4. Beynon MJ, Peel MJ (2001) Variable precision rough set theory and data discretisation: an application to corporate failure prediction. Omega 29(6):561–576
5. Carter C, Catlett J (1987) Assessing credit card applications using machine learning. IEEE Expert 2(3):71–79
6. Chen M, Ma L, Gao Y (2010) Vehicle detection segmentation based on adaboost and grabcut. In: IEEE International Conference on Progress in Informatics and Computing, pp 896–900
7. Chen MC, Huang SH (2003) Credit scoring and rejected instances reassigning through evolutionary computation techniques. Expert Syst Appl 24(4):433–441
8. Desai VS, Crook JN, Overstreet J (1996) A comparison of neural networks and linear scoring models in the credit union environment. Oper Res 95(2):24–37
9. Essa EM, Tolba AS, Elmougy S (2008) A comparison of combined classifier architectures for arabic speech recognition. In: International Conference on Computer Engineering & Systems, pp 149–153
10. Freund Y, Schapire RE (1995) A decision-theoretic generalization of on-line learning and an application to boosting. In: European Conference on Computational Learning Theory, pp 119–139
11. Gestel TV, Baesens B et al (2003) A support vector machine approach to credit scoring. Banken Financiewezen 2:73–82
12. Henley WE, Hand DJ (1996) A k-nearest-neighbour classifier for assessing consumer credit risk. J Roy Stat Soc 45(1):77–95
13. Henley WE, Hand DJ (1997) Construction of a k-nearest-neighbour credit-scoring system. IMA J Manag Math 8(4):305–321
14. Khashman A (2009) A neural network model for credit risk evaluation. Int J Neural Syst 19(4):285–294
15. Li H, Sun J (2008) Ranking-order case-based reasoning for financial distress prediction. Knowl Based Syst 21(8):868–878
16. Li H, Sun J (2009) Gaussian case-based reasoning for business failure prediction with empirical data in China. Inf Sci 179(1):89–108
17. Li H, Sun J (2009) Predicting business failure using multiple case-based reasoning combined with support vector machine. Expert Syst Appl 36(6):10085–10096

18. Li H, Sun J (2010) Business failure prediction using hybrid2 case-based reasoning (h2cbr). Comput Oper Res 37(1):137–151
19. Li H, Sun J, Sun BL (2009) Financial distress prediction based on or-cbr in the principle of k-nearest neighbors. Expert Syst Appl 36(1):643–659
20. Malhotra R, Malhotra DK (2002) Differentiating between good credits and bad credits using neuro-fuzzy systems. Eur J Oper Res 136(1):190–211
21. Malhotra R, Malhotra DK (2003) Evaluating consumer loans using neural networks. Soc Sci Electron Publishing 31(2):83–96
22. Ong CS, Huang JJ, Tzeng GH (2005) Building credit scoring models using genetic programming. Expert Syst Appl 29(1):41–47
23. Piramuthu S, Piramuthu S (1999) Financial credit-risk evaluation with neural and neurofuzzy systems. Eur J Oper Res 112(2):310–321
24. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536
25. Schebesch KB, Stecking R (2005) Support vector machines for classifying and describing credit applicants: detecting typical and critical regions. J Oper Res Soc 56(9):1082–1088
26. Steenackers A, Goovaerts MJ (1989) A credit scoring model for personal loans. Insur Math Econ 8(1):31–34
27. Thomas LC, Edelman DB, Crook JN (2002) Credit scoring and its applications. SIAM, Philadelphia
28. Varetto F (1951) Genetic algorithms in the analysis of insolvency risk. University of Illinois Press, Champaign
29. Wang Y, Wang S, Lai KK (2006) A new fuzzy support vector machine to evaluate credit risk. IEEE Trans Fuzzy Syst 13(6):820–831
30. West D (2000) Neural network credit scoring models. Comput Oper Res 27(11):1131–1152
31. Yobas MB, Crook JN, Ross P (2000) Credit scoring using neural and evolutionary techniques. IMA J Manag Math 11(2):111–125
32. Yu L, Wang S, Lai KK (2008) Credit risk assessment with a multistage neural network ensemble learning approach. Expert Syst Appl 34(2):1434–1444
33. Yu L, Wang SY et al (2008) Designing a hybrid intelligent mining system for credit risk evaluation. Syst Sci Complex 21(5):527–539
34. Yu L, Wang S, Cao J (2009) A modified least squares support vector machine classifier with application to credit risk analysis. Int J Inf Technol Decis Making 08(4):697–710
35. Zhou L, Lai KK, Yu L (2009) Credit scoring using support vector machines with direct search for parameters selection. Soft Comput 13(2):149–155

Effects of Urban and Rural Residents' Behavior Differences in Sports and Leisure Activity: Application of the Theory of Planned Behavior and Structural Equation Modeling

Linling Zhang[1,2](✉)

[1] School of Business, Sichuan University, Chengdu 610065, People's Republic of China
1010519067@qq.com

[2] Leisure Sports Department, Chengdu Sport Institute, Chengdu 610041, People's Republic of China

Abstract. Because the urban and rural residents' sports and leisure behaviors were different, to find out the factors influencing these differences was the purpose of this study. Therefore we selected urban and rural residents in Sichuan province as the research object to start our study. This study which was based on the Theory of Planned Behavior (TPB) used Structural Equation Modeling (SEM) as the research method. Firstly we constructed the conceptual model and put forward a series of assumptions. Then we used questionnaire survey to obtain first-hand information, at the same time we used ARMOS17.0 to work out the model. At last we verified our original assumption with the model. Research showed that: (1) The influence on the urban and rural residents' sports and leisure behavior from external objective factors (social atmosphere, venues, facilities, etc.) was much larger than that from the intrinsic objective factors (leisure time, income, physical condition, etc.); (2) The richness of sports and leisure activities, venues, facilities and the social atmosphere impacted the behavior of urban residents much more than that of rural residents; (3) The attitude to affirm and recognize sports and leisure had much more influence on urban residents; (4) While rural residents more cared about the views from family, friends even government.

Keywords: Sports and leisure activity · Urban and rural residents · Effects · Theory of planned behavior · Structural equation modeling supplier selection

1 Introduction

Under the Chinese specific dual social structure between urban and rural areas, social and economic activities, material conditions, living environment,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_7

life style, cultural consciousness between urban and rural areas appear a huge differentiation [5, 24]. In terms of sports and leisure activity, social system differences between the urban and rural result in gaps which come from capital investment, facilities construction, social organization development and concepts. And the leisure sports behavior of urban and rural residents also shows difference. What's more, with the accelerating of urbanization process, urban and rural demographic differences will also expand the difference of leisure sports behavior [11]. Therefore, the purpose of this study is trying to find out the main factors which cause huge difference and gap between urban and rural residents, and to think about how to close the gap.

Sports and leisure behavior is a dynamic process that people try to meet their leisure needs to participate in sports activities spontaneously and get satisfaction from it. The purpose that people take part in leisure sports activities mainly includes physical aspect (exercise, beauty fitness, self-defense, etc.) and emotional aspect (entertainment, relieve pressure, thrill-seeking, social communication, etc.) [19]. Western researches about sports and leisure behavior is mainly focused on the influence of behavior by using various theories. Trans-theoretical Model (TTM), Self-efficacy, Self-determination Theory (SDT), Leisure Constrains Theory, Theory of Planned Behaviors (TPB) are the representative theories [3, 4, 6–9, 15, 16, 18]. In 2006, Xu [21] who published a paper "The Study on the characteristics of the Residents' Leisure Sports Behaviors" in the Journal of Guangzhou Sport University opened the prelude of the domestic sports and leisure behavior research. Qiu [10, 12–14] who published many articles in the Chinese authoritative core journals is very outstanding in the study of sports and leisure behavior. It is worth mentioning Yu [23], Zhao [25], Ye [22] and others who introduced the theory of planned behavior (TPB) into the sports and leisure behavior study. Overall, the domestic scholars' research mainly focused on cities and towns and little focused on rural areas and it caused the lacking of empirical studies in such areas.

It is very hard to find out a clear answer from existing researches in the face of the question what is the deep reason that causes sports and leisure behavior differences between the urban and rural residents. In order to answer this question better, we introduce the Theory of Planned Behavior (TPB) and Structural Equation Modeling (SEM) as our theoretical and method guide and choose Sichuan province as the sample to try to reveal the main influencing factors of sports and leisure behavior of urban and rural residents and to analyze the relationship between these factors. We try to work out a theoretical analysis framework to explain the differences between urban and rural residents' sports and leisure behavior and this framework should play an important role to enrich and expand the leisure science research. According to the results, we put forward corresponding countermeasures and suggestions to improve the policy environment effectively and promote the development of urban and rural integration better.

2 Theory and Hypothesis

2.1 Introduction to the Theory of Planned Behavior

Theory of Planned Behaviors (TPB) is one of the most famous theories about the relationship between attitude and behavior in the field of social psychology. Based on the Theory of Reasoned Action (TRA), Ajzen [1] published "The Theory of Planned Behavior" in 1991, it marked the born of the formal formation of Planned Behavior Theory. In the TPB, the Behavior is affected by the Behavior Intention or the Perceived Behavioral Control. And the Behavior Intention is affected by the Attitude toward the Behaviors, Subjective Norm and Perceived Behavioral Control, at the same time Behavioral beliefs, Normative beliefs and Control beliefs respectively affect the Behavioral Attitude, Subjective Norm and Perceived Behavioral Control.

The effects of sports and leisure behavior involve economic factors (income, consumption idea), cultural factors (leisure concept, leisure atmosphere, leisure time, etc.), environmental factors (Sports venues, policy guarantee, product promotion, service level, etc.), etc. In this study we analyze problems from the perspective of the TPB and give out the main factors which affect the urban and rural residents' behavior. Below are the main factors.

(1) Attitude
If people keep positive attitudes on the sports and leisure activity, they will actively participate in the activity. It is obvious that the attitude can affect the behavior.

(2) Subjective Norm
It refers to whether the views and practices of people around will affect residents' participation in sports and leisure activities.

(3) Perceived Behavioral Control
It refers to whether the past experience and current conditions of residents have influence on their sports and leisure behaviors.

(4) Behavior Intention
It refers to if the residents are more willing to participate in sports and leisure activities subjectively, the probability that they actually participate in sports and leisure activities is greater. And the significant positive correlation between behavior and intention is also confirmed by many relative researches.

(5) Other Objective Factors
According to interview results from Sichuan residents, the main reasons that people do not attend sports and leisure activities include lack of enough time, lack of appropriate sports ground, lack of specific skills for their favorite sports, lack of information and lack of partner, etc.

2.2 Conceptual Model and Assumptions

According to the TPB and the previous investigation results, we build our conceptual model (Fig. 1). According to the key points in the model, we make the following theoretical assumptions:

Hypothesis 1 (H1): The attitude of sports and leisure has a significant positive influence on the behavior intention.

Hypothesis 2 (H2): The attitude of sports and leisure plays an intermediary role between subjective norm and behavior intention.

Hypothesis 3 (H3): The attitude of sports and leisure behavior plays an intermediary role between perceived behavioral control and behavior intention.

Hypothesis 4 (H4): Subjective norm has a significant positive influence on the behavior intention.

Hypothesis 5 (H5): The perceived behavioral control has a significant positive influence on the behavior intention.

Hypothesis 6 (H6): The peripheral objective factors have a significant positive influence on the TPB system.

Hypothesis 6a (H6a): The peripheral objective factor has a significant positive influence on the subjective norm.

Hypothesis 6b (H6b): The peripheral objective factors have a significant positive influence on the attitude.

Hypothesis 6c (H6c): The peripheral objective factors have a significant positive influence on the perceived behavioral control.

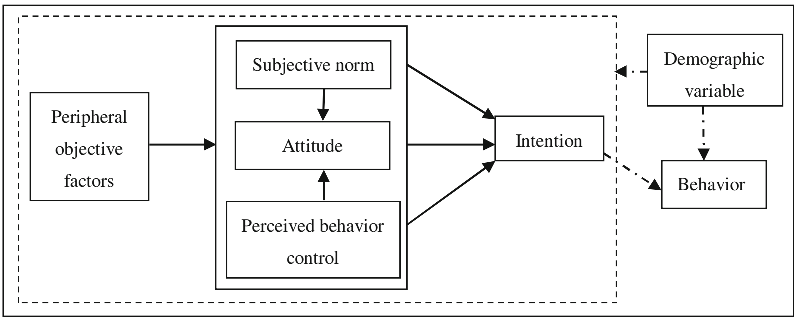


Fig. 1. The conceptual model of factors influencing sports and leisure behavior based on the TPB

3 Data Collection

3.1 Questionnaire Design

The influence factors of sports and leisure behavior mainly include the attitude, subjective norm, perceived behavioral control, peripheral objective factors and intention. The relevant information is collected through the questionnaire survey. This questionnaire chooses the Likert Scale, each question uses a measurement from strongly disagree (1 point) to strongly agree (5 points). All questions in Questionnaire come from the discussion of our team and the experts' suggestion. Below are the questions designed according to the conceptual model and the hypothesis (Table 1).

3.2 Data Collection

Empirical investigation in this study mainly involved Chengdu and other seventeen cities of Sichuan. We issued questionnaires two times. One is in spring of 2016 and the other is in summer of 2016.

For the first time, we mainly authorized students from Chengdu Sport Institute to issue the questionnaires in spring. During the vacation we issued 200 questionnaires through network and field investigation, and we received 62 effective questionnaires (43 ones from urban residents while only 19 ones from rural residents). Unfortunately, the effective recycling rate was very low and the reason was that: firstly, in the questionnaire we designed three filtering questions to filter bad ones; Secondly, the types of sports and leisure activities in rural areas are much less than these in urban areas, participation rate was also lower than that in town. In addition, in recent year rural migrant workers continued to increase in young adults, people left were only older and children. The older and children left were not able to finish questionnaire, so the amount of effective questionnaires was very small.

For the second time, we still authorized students from Chengdu Sport Institute to issue questionnaires in summer vacation. During the whole vacation we issued 400 questionnaires. This time we did some adjustments to issue more questionnaires to the rural residents. At last we recycled 168 effective questionnaires (83 ones from urban residents and 85 ones from rural residents).

To sum up, we issued 600 questionnaires out and recycled 230 effective ones at last. Among the effective questionnaires, 126 ones were from urban residents and the left 104 ones came from rural residents. Though sample size was not big enough but it could provide supports for this study.

3.3 Validity and Reliability Analysis

(1) Validity Analysis
Validity analysis of the questionnaire mainly uses the Factor Analysis method. We handle 20 questions in Table 1 as 20 variables which have influence on sports

Table 1. The questionnaire items assessing the TPB variables

Variables	Items (Question)
Attitude	a1. It is a pleasure to take part in sports and leisure activities
	a2. Sports and leisure activities are the best choice for entertainment and relaxation
	a3. Sports and leisure can help me to make more friends
	a4. Sports and leisure is an important part of my life
	a5. To participate in sports and leisure activities needs to insist on a long time, for which you are willing to give up some of the other things you like
Subjective norms	a6. The family supports me to participate in sports and leisure activities
	a7. Friends / colleagues / classmates support and appreciate me to participate in sports and leisure activities
	a8. When people see I take part in sports and leisure activities, they think I'm a man of life taste
	a9. The government encourages us to participate in sports and leisure activities
Perceived behavior control	a10. Enough time
	a11. Good income
	a12. Good physical condition
	a13. Enough information
	a14. It is very convenient to participate in sports and leisure activities
Peripheral objective factors	a15. The sports and leisure activities are abundant
	a16. Enough ground and facilities for sports and leisure activities
	a17. Good atmosphere of sports and leisure
Intention	a18. I am willing to participate in sports and leisure activities under the existing conditions
	a19. I will take part in sports and leisure activities in the future
	a20. I will persuade my family and friends together to participate in sports and leisure activities

and leisure behavior. By using SPSS19.0 software to calculate, Kaiser-Meyer-Olkin (KMO) is 0.871 and the significant probability of Bartlett Test of Sphericity statistic is 0.000 ($p < 0.01$), both of them indicate the data is suitable for factor analysis. Then we extract 5 common factors from the 20 variables, they can account for 63.876% of the total variation. Through rotary, we can see the factor-1 has a great influence on variables-a1, a2, a3, a4, a5, a18 and a19 (load value > 0.5); Factor-2 has a great influence on variables-a6, a7, a9 and a20; Factor-3 has a great influence on variables-a12, a15, a16 and a17; Factor-4 has a greater influence on variables-a10 and a13; and factor-5 has a great influence on variable-a11. While for variables a8 and a13, each of their load value on every common factor is less than 0.5, so we have reason to delete them. In the next step, we find that the item variable-a11 also should be removed through the tests for every common factor respectively.

(2) Reliability Analysis
We do the internal consistency test for the variables of every factor with Cronbach's alpha reliability analysis. And we find that 4 factors (attitude, subjective norm, periphery objective factors and intention) pass the test. If we delete the variables $a11$ and $a12$ in the factor-perceived behavioral control, the a-coefficient increases significantly to 0.698. In order to pass the internal consistency test, variables a11 and a12 should be deleted.

4 Modeling and Modification

4.1 Structural Equation Modeling

Structural Equation Modeling (SEM) is a common method in social science research. In the field of social science, the causal relationships among variables are complex sometimes, or these variables can't be observed directly, therefore traditional statistical methods can't solve this problem well. Since the 1980s, SEM has been developed quickly, it makes up for the deficiency of the traditional statistical method and becomes an important tool of multivariate data analysis [2,20]. SEM is a kind of multivariate analysis technology based on the regression. For it is of theory apriority, SEM is used to check the hypothesis between observed variables and latent variables, generally 5 samples are enough for every observed variable [17]. This research built the SEM of urban and rural residents' sports and leisure behavior influence factors with ARMOS17.0 software (Fig. 2).

4.2 The Urban Model

(1) Calculation and Correction
According to the SEM shown in Fig. 2, we take the 126 samples of the urban residents into the model. There are 5 paths that don't pass the significance test ($p < 0.01$) in the results and TLI index also fails in the inspection, so the model needs to be modified and optimized. After repeating the path correction, we delete the factors of perceived behavioral control, the results are obviously optimized (Tables 2 and 3). All indexes pass the test. Every variance estimation also passes the significance test ($p < 0.01$). In order to make the original model optimized better, we adopt Modification Index to fix the model again. We try

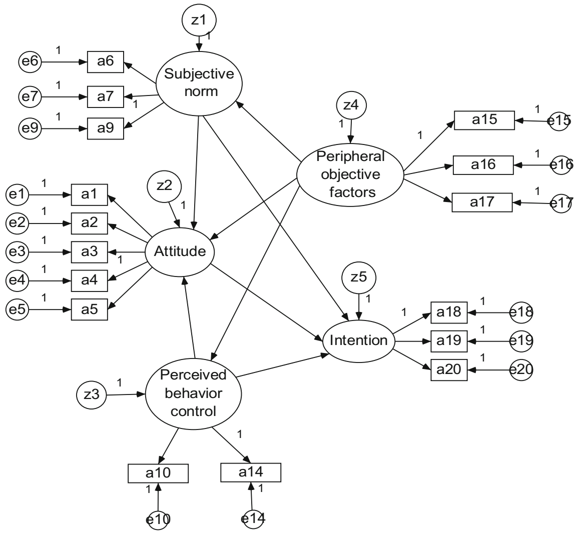


Fig. 2. Structural equation modeling

to modify the error terms of some observed variables by calculating MI value and find that it can optimize the index, but it can't improve the P value. On balance, we give up Modification Index and just keep the first adjusted result.

(2) Output
For the urban residents' sports and leisure behavior influence factor, we summarize the verification results of hypothesis 1 to 6 (Table 4). Then we can build out the SEM for urban residents.

4.3 The Rural Model

(1) Calculation and correction
According to the SEM shown in Fig. 2, we take the 104 samples of the rural residents into the model. 4 paths in the results don't pass the significance test ($p < 0.01$), and the AIC and EVCI indexes also fail in the test. It needs to correct the model. After repeating path correction, we delete the factor of perceived behavioral control and use Modification Index to revise the error terms of a4, a5 and a18 in the model. The results are optimized, all indexes pass the test, and every variance estimation also passed the significance test ($p < 0.01$). (Tables 5 and 6)

(2) Output
In view of the sports and leisure behavior influence factors of rural residents, we summarize the verification results of hypothesis 1 to 6 (Table 7). Then we can build out the SEM for rural residents.

Table 2. The adjusted coefficient estimates

			Estimate	S.E	C.R	P	Label	Estimate
Attitude	←	Peripheral objective factors	0.884	0.166	5.34	***	par_12	0.764
Subjective norm	←	Peripheral objective factors	0.327	0.099	3.294	***	par_13	0.433
Intention	←	Subjective norm	0.167	0.095	1.756	0.079	par_10	0.135
Intention	←	Attitude	0.648	0.078	8.258	***	par_11	0.807
a9	←	Subjective norm	1					0.564
a7	←	Subjective norm	1.548	0.248	6.251	***	par_1	0.867
a6	←	Subjective norm	1.581	0.253	6.259	***	par_2	0.86
a4	←	Attitude	1					0.831
a3	←	Attitude	0.836	0.128	6.528	***	par_3	0.57
a2	←	Attitude	0.81	0.089	9.069	***	par_4	0.743
a1	←	Attitude	0.672	0.079	8.549	***	par_5	0.71
a15	←	Peripheral objective factors	1					0.686
a16	←	Peripheral objective factors	0.859	0.187	4.581	***	par_6	0.514
a17	←	Peripheral objective factors	0.841	0.166	5.072	***	par_7	0.586
a18	←	Intention	1					0.843
a19	←	Intention	1.109	0.101	11.029	***	par_8	0.856
a20	←	Intention	1.113	0.124	8.976	***	par_9	0.73
a5	←	Attitude	0.91	0.107	8.485	***	par_14	0.706

*** $p < 0.01$

Table 3. The calculation results of the revised fitting index-2

Fit index	χ^2(df)	CMIN/DF	P	TLI	IFI	CFI	RMSEA	AIC	EVCI
Results	130.356 (73)	1.786	0	0.909	0.929	0.927	0.079	194.356	1.555
Conclusion	Suitable	Pass	Pass	Pass	Pass	Pass	Suitable	Pass	Pass

Table 4. Hypothesis testing results of the effects of urban residents' behavior in sports and leisure

Hypotheses	Validation results
H1: The Attitude of sports and leisure has a significant positive influence on the behavior intention	Support
H2: The attitude of sports and leisure plays an intermediary role between subjective norm and behavior intention	Nonsupport
H3: The attitude of sports and leisure behavior plays an intermediary role between perceived behavioral control and behavior intention	Nonsupport
H4: Subjective norm has a significant positive influence on the behavior intention	Support
H5: The perceived behavioral control has a significant positive influence on the behavior intention	Nonsupport
H6: The peripheral objective factors have a significant positive influence on the TPB system	Partial support
H6a: The peripheral objective factor has a significant positive influence on the subjective norm	Support
H6b: The peripheral objective factors have a significant positive influence on the attitude	Support
H6c: The peripheral objective factors have a significant positive influence on the perceived behavioral control	Nonsupport

Table 5. The adjusted coefficient estimates

			Estimate	S.E	C.R	P	Label	Estimate
Subjective norm	←	Peripheral objective factors	0.263	0.072	3.633	***	par_13	0.382
Attitude	←	Peripheral objective factors	0.69	0.132	5.245	***	par_12	0.582
Attitude	←	Subjective norm	0.444	0.141	3.14	***	par_15	0.257
Intention	←	Subjective norm	0.401	0.098	4.094	***	par_10	0.289
Intention	←	Attitude	0.504	0.058	8.642	***	par_11	0.628
a9	←	Subjective norm	1					0.49
a7	←	Subjective norm	1.834	0.243	7.563	***	par_1	0.911
a6	←	Subjective norm	1.76	0.232	7.572	***	par_2	0.876
a4	←	Attitude	1					0.799
a3	←	Attitude	0.85	0.093	9.171	***	par_3	0.604
a2	←	Attitude	0.892	0.071	12.55	***	par_4	0.791
a1	←	Attitude	0.778	0.061	12.86	***	par_5	0.809
a15	←	Peripheral objective factors	1					0.698

Table 5. (*Continued*)

			Estimate	S.E	C.R	P	Label	Estimate
a16	←	Peripheral objective factors	0.769	0.144	5.341	***	par_6	0.469
a17	←	Peripheral objective factors	0.839	0.137	6.129	***	par_7	0.575
a18	←	Intention	1					0.868
a19	←	Intention	1.14	0.067	17.094	***	par_8	0.892
a20	←	Intention	1.116	0.083	13.469	***	par_9	0.755
a5	←	Attitude	0.777	0.069	11.333	***	par_14	0.61

*** $p < 0.01$

Table 6. The calculation results of the revised fitting index-2

Fit index	χ^2(df)	CMIN/DF	P	TLI	IFI	CFI	RMSEA	AIC	EVCI
Results	133.622(70)	1.909	0	0.947	0.935	0.96	0.063	203.622	0.889
Conclusion	Suitable	Pass	Pass	Pass	Pass	Pass	Suitable	Pass	Pass

Table 7. Hypothesis testing results of the effects of rural residents' behavior in sports and leisure

Hypotheses	Validation results
H1: The Attitude of sports and leisure has a significant positive influence on the behavior intention	Support
H2: The attitude of sports and leisure plays an intermediary role between subjective norm and behavior intention	Support
H3: The attitude of sports and leisure behavior plays an intermediary role between perceived behavioral control and behavior intention	Nonsupport
H4: Subjective norm has a significant positive influence on the behavior intention	Support
H5: The perceived behavioral control has a significant positive influence on the behavior intention	Nonsupport
H6: The peripheral objective factors have a significant positive influence on the TPB system	Partial support
H6a: The peripheral objective factor has a significant positive influence on the subjective norm	Support
H6b: The peripheral objective factors have a significant positive influence on the attitude	Support
H6c: The peripheral objective factors have a significant positive influence on the perceived behavioral control	Nonsupport

5 Results

In view of the above empirical analysis, there are some differences between urban residents and rural residents in sports and leisure behavior influence factor.

From the perspective of direct effects (Table 8), the influence of peripheral objective factor that affected urban residents' sports and leisure attitude was larger 0.182 units than what affected rural residents, the influence of subjective norm on urban residents was larger 0.051 units than that on rural residents. For the influence of sports and leisure attitude on the intention, urban residents were higher 0.179 units than rural residents. The subjective norm only affected the intention of urban residents but did not affect their attitude, however subjective norm not only affected the intention of rural residents but also affected their attitude. The subjective norm had more influences on the intention of rural residents.

From the perspective of indirect effects (Table 8), the influence of peripheral objective factors on the intention of urban residents was 0.675 units, it was higher than that on rural residents. At the same time, the peripheral objective factors also indirectly affected the attitude of the rural residents. The subjective norm could indirectly affect the intention of rural residents through attitude, but it could not affect urban residents.

Table 8. Direct and indirect effect (Standardized results)

	Urban				Rural			
	Direct effect/Indirect effect				Direct effect/Indirect effect			
	Peripheral objective factors	Attitude	Subjective norm	Intention factors	Peripheral objective	Subjective norm	Attitude	Intention
Attitude	.764/	-	-	-	.582/.098	.257/	-	-
Subjective norm	.433/	-	-	-	.382/	-	-	-
Intention	/.675	.807/	.135/	-	/.537	.289/.162	.628/	-
a5	/.540	.706/	-	-	/.415	/.157	.610/	
a20	/.493	/.588	/.099	.730/	/.406	/.341	/.474	.755/
a19	/.578	/.690	/.116	.856/	/.480	/.402	/.560	.892/
a18	/.569	/.680	/.114	.843/	/.467	/.392	/.545	.868/
a17	.586/	-	-	-	.575/	-	-	-
a16	.514/	-	-	-	.469/	-	-	-
a15	.686/	-	-	-	.698/	-	-	-
a1	/.543	.710/	-	-	/.550	/.208	.809/	-
a2	/.568	.743/	-	-	/.538	/.204	.791/	-
a3	/.436	.570/	-	-	/.411	/.156	.604/	-
a4	/.635	.831/	-	-	/.544	/.206	.799/	-
a6	/.373	-	.860/	-	/.335	.876/	-	-
a7	/.376	-	.867/	-	/.348	.911/	-	-
a9	/.245	-	.564/	-	/.187	.490/	-	-

From the perspective of total effects (Table 8), the influence of peripheral objective factors on the attitude, subjective norm and intention of urban residents was greater than that of rural residents. For the influence of attitude on the intention, it was larger for urban residents than for rural residents. For the influence of subjective norm on the intention, it was larger for rural residents than for urban residents. In addition, the subjective norm also affected the attitude of rural residents, but it did not affect the attitude of urban residents.

6 Discussion and Conclusion

The sports and leisure behavior of urban residents was affected by the intention, attitude, subjective norm and peripheral objective factors. The Subjective norm, attitude and intention were the intermediary variables. The subjective norm and the behavior attitude were two independent variables, besides, they were both affected by the peripheral objective factors, at the same time they indirectly affected the sports and leisure behavior by impacting behavioral intention. Sports leisure and behavior of rural residents was also affected by the intention, attitude, subjective norm and peripheral objective factors. The subjective norm, behavior, attitude and intention were the intermediary variables. The relationships among them were more complicated than the relationships of urban residents. The peripheral objective factors affected the subjective norm and attitude, at the same time they affected the sports and leisure behavior through behavioral intention. In addition, the subjective norm affected the attitude of rural residents, but it did not affect the attitude of urban residents.

After we calculated the path coefficient, we got a point of view about the total effect. The influence of peripheral objective factors on the attitude, subjective norm and intention of urban residents was higher than that of rural residents. For the influence of attitude on the intention, it was higher for urban residents than for rural residents. For the influence of subjective norm on the intention, it was higher for rural residents than for urban residents. That is to say, the influence of external objective factors (social atmosphere, sports ground, facilities, etc.) on the sports and leisure behavior of both urban and rural residents was much larger than the influence of intrinsic objective factors (personal leisure time, personal income, physical condition, etc.). The diversity and richness of sports and leisure activities, sports ground, facilities, and social atmosphere impacted urban residents more than rural residents. Admissive and positive attitude toward sports and leisure activities influenced urban residents more. While rural residents more cared about the view of their family, friends and even the government about their participation in sports and leisure activities.

Acknowledgement. We thank Guo Xinyan and Liu Ying for technical assistance; Zhang Xu for guidance. This work was supported by Sichuan Social Science Fund (Grant # SC16XK025).

References

1. Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process 50:179–211
2. Ajzen I (2016) Ownership structure, diversification, and corporate performance based on structural equation modeling. J Real Estate Portfolio Manage 22:63–73
3. Beville JM, Meyer M et al (2014) Gender differences in college leisure time physical activity: application of the theory of planned behavior and integrated behavioral model. J Am Coll Health 62(3):173–184
4. Boudreau F, Godin G (2014) Participation in regular leisure-time physical activity among individuals with type 2 diabetes not meeting canadian guidelines: the influence of intention, perceived behavioral control, and moral norm. Int J Behav Med 21:918–926 (in Chinese)
5. Caldwell JT, Ford CL et al (2016) Intersection of living in a rural versus urban area and race/ethnicity in explaining access to health care in the United States. Am J Public Health 106(8):1463–1469
6. Cerin E, Vandelanotte C et al (2008) Recreational facilities and leisure-time physical activity: an analysis of moderators and self-efficacy as a mediator. Health Psychol 27(2):126–135
7. Heesch KC, Masse LC (2004) Do women have enough time for leisure-time physical activity. Res Q Exerc Sport 75(1):A101–A101
8. Lloyd K, Little DE (2010) Self-determination theory as a framework for understanding women's psychological well-being outcomes from leisure-time physical activity. Leisure Sci 32(4):369–385
9. Nurmi J, Hagger MS et al (2016) Relations between autonomous motivation and leisure-time physical activity participation: the mediating role of self-regulation techniques. J Sport Exerc Psychol 38(2):128–137
10. Qiu Y (2008) Research on restricted factors of leisure sport behavior developing stagełan assumption theory framework. China Sport Sci 28:71–75 (in Chinese)
11. Qiu Y (2009) Research on the Motivations and Constraints of the Stages of Leisure Physical Activity. Zhejiang University Press (in Chinese)
12. Qiu Y (2011) Research on the negotiation strategies of the leisure sports activities. China Sport Sci 31:8–16 (in Chinese)
13. Qiu YJ, Jiao XU (2014) Research on the characteristics of constraints to women's leisure sports activities and the relationship with behaviors. China Sport Sci 32:25–33 (in Chinese)
14. Qiu YJ, Liang MY (2012) Qualitative research on constraints of Chinese women leisure sports behavior-based on the perspective of the theory of social gender. China Sport Sci 34:75–82 (in Chinese)
15. Sabiston CM, Crocker PRE (2008) Exploring self-perceptions and social influences as correlates of adolescent leisure-time physical activity. J Sport Exerc Psychol 30(1):3
16. Shields M, Ellingson L et al (2008) Determinants of leisure time physical activity participation among Latina women. Leisure Sci 30(5):429–447
17. Simpson VL, Hyner GC, Anderson JG (2013) Lifestyle behavior change and repeat health risk appraisal participation: a structural equation modeling approach. Am J Health Promot 28(2):128–135
18. Stanis SAW, Schneider IE, Russell KC (2009) Leisure time physical activity of park visitors: retesting constraint models in adoption and maintenance stages. Leisure Sci 31(3):287–304

19. VanSickle J, Schaumleffel NA (2016) Developing recreation, leisure, and sport professional competencies through practitioner/academic service engagement partnerships. Schole A J Leisure Stud Recreation Educ 31(2):37–55
20. Vanwesenbeeck I, Walrave M, Ponnet K (2016) Young adolescents and advertising on social network games: a structural equation model of perceived parental media mediation, advertising literacy, and behavioral intention. J Advertising 45(2):1–15
21. Xu J (2006) The study on the characteristics of the residents' leisure sports behaviors. J Guangzhou Sport Univ 26:97–100 (in Chinese)
22. Ye W (2014) Theory and empirical research on the intention of leisure sports behavior. Sport Sci Technol 35:102–104
23. Xu Y (2013) A study on the citizens' behavior intention of sports leisure based on the theory of planned behavior. Yunnan Geogr Environ Res 25:71–76 (in Chinese)
24. Zhao B (2011) Consumer ethical beliefs associated with its birthplace: empirical research of China's urban and rural dual society background. Manage World 1:92–100 (in Chinese)
25. Zhao Q, Xu N et al (2014) The study of the relation between leisure sports consumption culture and consumption behaviors. Sichuan Sports Sci 33:101–106 (in Chinese)

Fast Multiobjective Hybrid Evolutionary Algorithm Based on Mixed Sampling Strategy

Wenqiang Zhang[1,2(✉)], Yu Wang[1,2], Chunxiao Wang[1,2], Le Xiao[1,2], and Mitsuo Gen[3]

[1] Henan University of Technology, Zhengzhou, China
[2] Key Laboratory of Grain Information Processing and Control, Ministry of Education, Zhengzhou, People's Republic of China
zhangwq@haut.edu.cn
[3] Fuzzy Logic Systems Institute, Tokyo University of Science, Tokyo, Japan

Abstract. In this paper, a fast multiobjective hybrid evolutionary algorithm (MOHEA) is proposed to solve the multiobjective optimization problem (MOOP) in achieving a balance between convergence and distribution with computational complexity. The proposed algorithm, MOHEA, improves the vector evaluated genetic algorithm (VEGA) by combing a new sampling strategy according to the Pareto dominating and dominated relationship-based fitness function. VEGA is good at searching the edge region of the Pareto front, but it has neglected the central area of the Pareto front, and the new sampling strategy prefers the center region of the Pareto front. The mixed sampling strategy improves the convergence performance and the distribution performance while reducing the computational time. Simulation experiments on multiobjective test problems show that, compared with NSGA-II and SPEA2, the fast multiobjective hybrid evolutionary algorithm is better in the two aspects of convergence and distribution, and has obvious advantages in the efficiency.

Keywords: Evolutionary algorithm · Mixed sampling · Multi-objective optimization · Styling · Benchmark

1 Introduction

Multiobjective evolutionary algorithms (MOEAs) have been identified as very suitable for solving multiobjective optimization problems (MOP) [4]. MOEA is a heuristic method, and is useful for searching the Pareto optimal solutions of a MOP by global and local search between generations. From the results, MOEA can provide decision maker (DM) with a number of more practical solutions. So the DM can make decisions based on actual needs. In the optimization process of a single objective, the collaboration of other objectives should be considered at the same time. Meanwhile, maintaining the trend of simultaneous evolution

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_8

make the Pareto front solutions as close as possible. Simultaneously, the real Pareto optimal set is evenly distributed.

Moreover, a number of evolutionary algorithms have their own characteristics. Each algorithm is effective to solve the problem of limited areas of application. No algorithm can solve all the problems. How to mix these evolutionary algorithms to play their respective strengths has become one of the challenging research fields in evolutionary algorithms.

In order to compensate for the shortcomings of single MOEA, more researchers began to combine different algorithm to generate hybrid algorithm. In 2012, Zhang and Fujimura proposed an improved vector evaluation of genetic algorithm with archive (IVEGA-A) [6]. The IVEGA-A combines the strong convergence ability of the VEGA [7] method to the Pareto frontier boundary region [11] and the elite population updating mechanism based on the new fitness function to ensure the overall performance of the algorithm in the central region of the Pareto front. Zhang and Li introduced differential evolution strategy into MOEA/D [5].

In this paper, we propose a multi-objective hybrid evolutionary algorithm (MOHEA) to solve multi-objective optimization problem for improving the convergence, the distribution and reducing the computational time. The rest of the paper is organized as follows: An overview of sampling strategy is given in Sect. 2. MOHEA is summarized in Sect. 3. Section 4 presents experimental results to illustrate the efficiency of the algorithm.

2 Mixed Sampling Strategy

2.1 Pareto Dominating and Dominated Relationship-Based Fitness Function

In this paper, a new fitness function based on Pareto dominating and dominated relationship-based fitness function (PDDR-FF) is used to evaluate the individual. The fitness function value of an individual S_i is calculated by the following function:

$$\text{eval}(S_i) = q(S_i) + \frac{1}{p(S_i) + 1},\ i = 1, 2, \cdots, \text{popSize}, \tag{1}$$

where $q(S_i)$ is the number of individuals that dominate S_i, $p(S_i)$ is the number of individuals that dominated by S_i. popSize is the size of the population. If the individual S_i is non-dominated solution, $q(S_i) = 0$. Therefore, the fitness function of the non-dominated individual S_i is related to $p(S_i)$. The more individuals are dominated by S_i, the smaller the value of $1/(p(S_i) + 1)$, and the smaller the value of PDDR-FF. If $q(S_i) = 0$, $p(S_i) = 0$, the PDDR-FF value of S_i is 1. It means there is no individual dominates S_i and no individual is dominated by S_i. Therefore, the PDDR-FF value of the non-dominated solution will no more than 1. S_i is dominated individual, $q(S_i) \geq 1$; the less individuals that dominate S_i, the smaller $q(S_i)$. In addition, the more individuals that are dominated by

S_i, the bigger $p(S_i)$ and the smaller $1/(p(S_i)+1)$. Then, the value of PDDR-FF is smaller.

The smaller value of PDDR-FF means that the more individuals dominated by S_i and the less individuals dominating S_i. Therefore, the smaller the PDDR-FF value, the better. And besides, PDDR-FF can clearly identify the dominant and non-dominated solutions. If an individual is a non-dominant, the fitness function value is not more than 1. Even if the individuals are all non-dominated, the individuals with different number of dominant are endowed with different fitness function values. It is obvious that the fitness function value (close to 0) of the non-dominated individuals locating near the central area of the Pareto front having big domination area is smaller than the value (close to 1) of the individual in the edge region.

2.2 Mixed Sample Strategy of VEGA and PDDR-FF

The sampling strategy of PDDR-FF has the advantage of being able to converge to the central region of Pareto frontier. But the convergence of the edge region is not so good. Only using the PDDR-FF can cause the uneven distribution performance. Compared to the central area of Pareto frontier, the sampling strategy of VEGA is more preferred than the edge region, and the single use of VEGA sampling can also lead to the problem of the distribution performance. Therefore, the two sampling strategy can be combined to improve the overall performance of the algorithm and reduce the computation time.

3 Fast Multi-Objective Hybrid Evolutionary Algorithm

3.1 Detailed Process of MOHEA

The strong convergence capability of mixed sample strategy ensures that the MOHEA has ability to converge to the center and the edge region of true Pareto frontier. Meanwhile, different convergence regions ensure the distribution.

The specific algorithm flow of MOHEA is shown in Fig. 1.

The specific process of MOHEA is as follows:

Step 1. Initialization of algorithm parameters: The size of the population: popSize, the size of archive:archiveSize, the maximum number of calculation:maxEvaluations.

Step 2. Creating initial population $P(t)$ and initial archive $A(t)$.

Step 3. Loop $d)-g)$ step until the termination condition is satisfied: The number of the calculation achieves the maximum.

Step 4. Selection (VEGA sampling strategy) Using the VEGA sampling strategy to select good individuals as a part of the mating pool. For the two objectives of problem, at first all individuals in the population are ordered according to the value of object function 1 and select $\left|\frac{P(t)}{2}\right|$ individuals with replacement into the sub-population 1 in order. In a similar way, considering objective 2 and select $\left|\frac{P(t)}{2}\right|$ individuals with replacement into the sub-population 2.

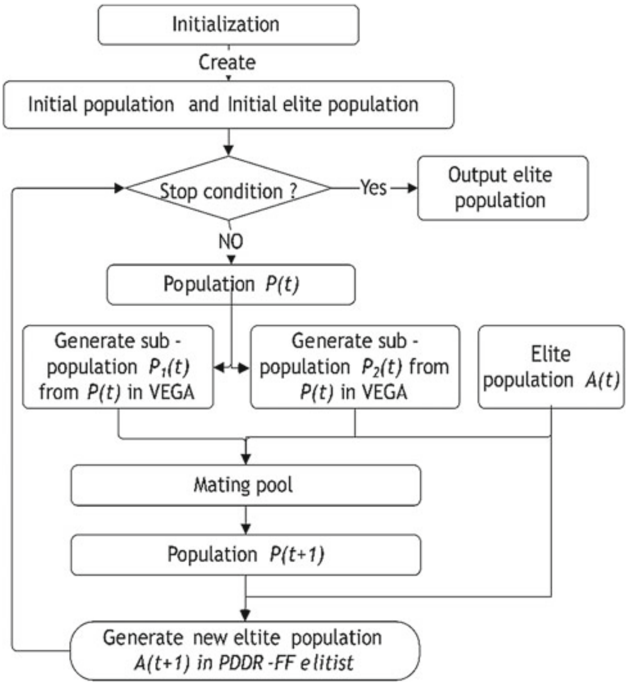


Fig. 1. The specific algorithm flow of MOHEA

Step 5. Generation of the mating pool: mixing the sub-populations and $A(t)$ to form the mating pool. In the mating pool, sub-population 1 stores the excellent individuals for the objective 1 and sub-population 2 keeps the excellent individuals for the another objective. The individuals with good PDDR-FF value are stored in $A(t)$. For the two objectives optimization problem, the size of these two sub populations and archive are set to the half size of the population. As a result, there are three parts in the mating pool, 1/3 of the individuals serve the objective 1, 1/3 of the individuals tend to the objective 2 and the remaining individuals obeys both the two objectives.

Step 6. Genetic operations: according to the selection operator, two individuals are selected from the mating pool as the parent and the new generation $P(t+1)$ is produced after genetic operations (crossing and mutation).

Step 7. The updating of the archive (PDDR-FF sampling strategy): mixing the new generation P(t+1) and the archive to form the temp population $A'(t)$. Calculate the fitness function value of all individuals in the population, according to the ascending order and individuals with the smallest of the $\mid A(t) \mid$ values are chosen as $A(t+1)$.

The overall procedure of MOHEA is described as the follows Fig. 2.

```
procedure: MOHEA
input: data set of problem, EA parameters
output: Pareto optimal solutions E
begin
    t ← 0;
    initialize P(t) by random key-based encoding routine;
    calculate two objectives by decoding routine;
    calculate fitness eval(P) by PDDR-FF approach;
    create Archive A(t) by PDDR-FF based updating routine;
    create Pareto E(t) by nondominated routine;
    while (not terminating condition) do
        create mating pool M(t) from P(t) by BTS routine;
        combine mating pool M(t) and Archive A(t) as M(t);
        create P(t+1) from M(t) by crossover routine;
        create P(t+1) from M(t) by mutation routine;
        calculate objectives by decoding routine;
        calculate fitness eval(P+1) by PDDR-FF approach;
        update Archive A(t+1) by PDDR-FF routine;
        update Pareto E(t) by nondominated routine;
        t ← t + 1;
    end
    output Pareto optimal solutions E(t)
end;
```

Fig. 2. The overall procedure of MOHEA

3.2 Genetic Encoding

The genetic code operating is real-coded in this paper. Each gene of the chromosome is consists of a random, evenly distributed, floating-point number in $[0, 1]$, or the true value of the decision variable. Each chromosome is a vector of a floating point type and the length of the chromosome is determined by the number of decision variables. Real coding makes the complex genetic operation easy to implement. It enhances the accuracy of the operation and strengthens the search ability of the algorithm to the large space.

For example, the number of benchmark test problem ZDT6 [10] decision variables is 10. Suppose vector $V = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1)$ is a set of ZDT6 solution, the vector V can be directly used as the chromosome corresponding gene coding sequence. Or the vector V' may be used as the gene sequence of the chromosome according to the mapping relationship in the Fig. 3.

3.3 Crossover Operator

The crossover operator used in this section is the Simulated Binary Crossover (SBX) proposed by Deb [1]. When the real-coded chromosomes are crossed by the SBX, the distribution parameter β of the parents need to be calculated by the equation:

$$\beta = \begin{cases} (2u)^{\frac{1}{\eta+1}}, & u \leq 0.5 \\ \frac{1}{2(1-u)}^{\frac{1}{\eta+1}}, & \text{othersize}, \end{cases} \tag{2}$$

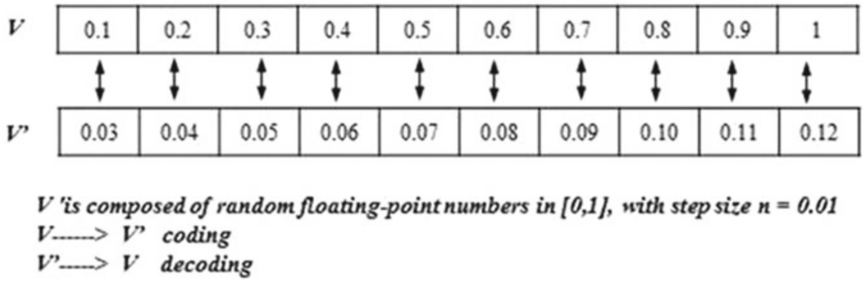


Fig. 3. Real-Coded mapping

where $u \in U(0,1)$, η represents the cross-distribution index and be determined by the decision maker. The larger the value, the smaller the distance between the progeny individual and the parents. Conversely, the greater the distance. Parent S_1 and S_2 can produce offspring s_1 and s_2 by the following equation:

$$\begin{cases} S_1 = 0.5\left[(1+\beta)\,S_1 + (1-\beta)\,S_2\right] \\ S_2 = 0.5\left[(1+\beta)\,S_1 + (1-\beta)\,S_2\right]. \end{cases} \tag{3}$$

3.4 Crossover Operator

The mutation operator used in this section is a polynomial variation proposed by Deb and Goyal [2]. The main forms of variation are:

$$S' = S + \delta\omega_{\max}, \tag{4}$$

where S is a normal individual and S' is a variant individual. ω_{max} is the maximum variable value of decision variables. δ is calculated as:

$$\delta = \begin{cases} (2v)^{\frac{1}{\gamma+1}} - 1, & v < 0.5 \\ 1 - [2(1-v)]^{\frac{1}{\gamma+1}}, & v \geq 0.5, \end{cases} \tag{5}$$

where $v \in V(0,1)$. $\gamma > 0$ represents the variance distribution index and be determined by the decision maker.

3.5 Selection Operator

In this study, NSGA-II [3] and SPEA2 [13] use binary contention selection as the selection operator. SPEA2 need to consider the dominance of the individual relationship. In this context, NSGA-II also needs to note the crowding distance of individuals. The mating pool of MOHEA consists of two parts: one is the individuals that are excellent for a single target. In the selection process, the individuals in the current population are first sorted according to the number of targets and the sub-population with excellent single-target value is selected in order. All sub-populations are included as part of the mating pool. The other is the non-dominated individuals in the elite population. In order to ensure sufficient information exchange between individuals in the two populations, it is necessary to randomly select two individuals from the mating pool during genetic manipulation.

4 Experiments and Discussion

The benchmark problem is used to test the MOHEA proposed in this paper. At the same time, two classical multi-objective evolutionary algorithms, NSGA-II and SPEA2, are tested and compared under the same conditions. The test function uses the ZDT and DTLZ problems that are currently used in the MOEAs domain. The algorithm is implemented in Java language. The hardware condition of the PC computer is: CPU Core 2, clock speed 2.8G, memory 2G, and operating system Windows XP.

4.1 Evaluation Index

Let S_j be the solutions of each algorithm ($j = 1, 2, 3$). In this study, the test considers two kinds of common convergence and distribution performance measures.

Coverage $C(S_1, S_2)$ is the percent of individuals dominated by S_1 in S_2 [12]. The bigger the $C(S_1, S_2)$ value, the better the convergence performance of S_1.

Spacing SP is the standard deviation of the closest distance of individuals in S_j [8]. The smaller value of SP means the better distribution performance.

The mathematically defined expressions for the 12 benchmark tests used in the experiment are shown in [10]. Among them, ZDT1, ZDT2 and ZDT3 have 30 decision variables, ZDT4 and ZDT6 have 10 decision variables. The decision variables and the number of targets of DTLZ problem can be expanded arbitrarily. For k and $\mid xk \mid$ values, if the target number is 2, then $k = 2$.

The parameters used in the three algorithms are as same as the original. The detailed values are shown in the following.

- Population size: 100
- Elite population size: 100
- Crossover rate: 0.4
- Mutation rate: 0.8
- Cross distribution index: 15
- Variation distribution index: 20
- Max generation: 500

The three algorithms have the same crossover and mutation operator. Fitness evaluation mechanism and the elite strategy is not the same. It should be pointed out that NSGA-II and SPEA2 need to select all non-dominated solutions from the elite as Pareto sets. MOHEA updates the Pareto set by choosing non-dominated solutions among the elite and the population. For the optimal stopping criterion of the algorithm, this paper uses the most widely used function evaluation number as the termination condition of the algorithm.

The performance evaluation index of the algorithm adopts the convergence index C and the distributed index SP. The results of the three algorithms running 30 times are compared. The time of each algorithm running is recorded. Thus the algorithm efficiency index CPU is obtained. A boxplot was used to evaluate the alignment results of the algorithm during 30 calculation results [9]. Box diagram is an important tool for statistical analysis. The image can clearly reflect the distribution of data. It is the most effective way to display EAs graphically. The upper line of the box represents the quartile line of the sample. The lower line indicates the lower quartile line of the sample. The middle line represents the bisector (mean). The top of the box represents the maximum value of the sample and the bottom is the minimum value. The gap in the middle of the box represents the confidence interval, and the other scatter points represent the outliers.

4.2 Discussion of the Results

The evaluation indexes C, SP and CPU obtained from MOHEA, NSGA-ii, and SPEA2 are shown in Fig. 3, and the average and variance of each index are shown in Tables 1 and 3. It can be concluded that MOHEA performs better on most problems than NSGA-II and SPEA2 in terms of convergence. On questions ZDT1, ZDT2, ZDT6 and DTLZ7, NSGA-II and SPEA2 performed well with MOHEA as shown in Table 1. For the distribution, it can be seen from Tables 2 and 3 that the distribution of MOHEA is better than the other two algorithms except for the problems ZDT1 and DTLZ6. Since the number of variables and the range of variables make the focus of the 12 test functions different, it is difficult for the same algorithm to perform optimally on all test functions.

The efficiency of the algorithm can be discussed from two aspects. M is the number of targets and N mean the population number. In terms of distance com-

Table 1. C index

Questions	C(NSGA-II, MOHEA)	C(MOHEA, NSGA-II)	C(SPEA2, MOHEA)	C(MOHEA, SPEA2)
ZDT1	7.91e-011.5e-01	1.16e-011.4e-01	4.20e-011.8e-01	4.77e-012.0e-01
ZDT2	6.26e-027.1e-02	8.31e-016.8e-02	4.52e-012.6e-01	4.22e-012.2e-01
ZDT3	1.54e-011.5e-01	6.39e-012.2e-01	2.95e-012.2e-01	4.29e-011.9e-01
ZDT4	1.79e-013.2e-01	4.56e-013.1e-01	2.74e-013.2e-01	2.82e-012.6e-01
ZDT6	8.18e-011.9e-01	9.66e-022.1e-01	5.06e-012.6e-01	3.85e-012.6e-01
DTLZ1	1.98e-012.2e-01	6.04e-012.9e-01	3.06e-012.5e-01	4.47e-012.7e-01
DTLZ2	1.60e-022.9e-02	1.01e-012.1e-02	6.00e-031.0e-02	9.40e-022.2e-02
DTLZ3	1.79e-011.7e-01	4.24e-012.5e-01	3.25e-012.3e-01	4.43e-012.3e-01
DTLZ4	2.20e-021.9e-02	9.70e-022.1e-02	2.40e-022.1e-02	9.70e-022.5e-02
DTLZ5	2.50e-022.7e-02	7.90e-021.6e-02	3.00e-036.4e-03	7.40e-021.5e-02
DTLZ6	1.88e-011.3e-01	6.98e-011.8e-01	3.34e-012.9e-01	5.15e-012.7e-01
DTLZ7	7.64e-011.4e-01	1.51e-011.3e-01	5.15e-012.3e-01	3.99e-012.2e-01

Table 2. SP index

Questions	NSGA-II	SPEA2	MOHEA
ZDT1	1.12e-021.8e-03	1.12e-021.8e-03	1.85e-026.2e-03
ZDT2	2.03e-023.4e-03	2.03e-023.4e-03	1.18e-022.8e-03
ZDT3	1.06e-021.1e-03	1.06e-021.1e-03	9.33e-031.3e-03
ZDT4	1.27e+006.2e-01	1.27e+006.2e-01	3.20e-014.0e-01
ZDT6	5.16e-023.2e-02	5.16e-023.2e-02	2.11e-024.3e-03
DTLZ1	1.31e+004.1e-01	1.31e+004.1e-01	8.39e-011.7e+00
DTLZ2	4.46e-036.2e-04	4.46e-036.2e-04	4.01e-034.1e-03
DTLZ3	2.06e+013.0e+01	2.06e+013.0e+01	4.73e+001.7e+00
DTLZ4	6.54e-036.7e-04	6.54e-036.7e-04	4.45e-034.7e-04
DTLZ5	4.23e-033.6e-04	4.23e-033.6e-04	2.47e-031.7e-03
DTLZ6	9.05e-021.8e-02	9.05e-021.8e-02	1.01e-011.3e-02
DTLZ7	1.83e-027.1e-03	1.83e-027.1e-03	1.79e-022.9e-03

Table 3. CPU index

Questions	NSGA-II	SPEA2	MOHEA
ZDT1	7.92e+046.8e+03	1.29e+042.6e+03	8.29e+039.9e+02
ZDT2	1.03e+058.1e+03	1.20e+041.9e+03	7.80e+039.3e+02
ZDT3	7.28e+046.1e+03	1.19e+042.0e+03	9.11e+037.2e+02
ZDT4	6.80e+045.3e+03	1.19e+041.9e+03	9.06e+038.0e+02
ZDT6	8.92e+047.6e+03	1.25e+042.1e+03	8.73e+036.9e+02
DTLZ1	8.04e+048.8e+03	1.21e+041.9e+03	8.95e+037.6e+02
DTLZ2	1.06e+051.1e+04	1.16e+041.6e+03	6.88e+035.5e+02
DTLZ3	9.99e+041.8e+04	1.24e+041.9e+03	8.76e+037.9e+02
DTLZ4	1.25e+051.9e+04	1.32e+042.0e+03	8.77e+038.2e+02
DTLZ5	6.57e+043.2e+03	1.38e+041.3e+03	9.23e+039.8e+02
DTLZ6	7.99e+046.3e+03	1.22e+041.7e+03	8.73e+039.5e+02
DTLZ7	1.12e+052.0e+04	1.34e+041.9e+03	8.78e+031.1e+03

putation, the time complexity of NSGA-ii is $O\ (mN \log_2 N)$. The computational complexity of SPEA2 is $O\ (mN\ \log_2\mathrm{N})$. Since there is no distance mechanism in the MOHEA, the computational complexity of MOHEA is 0. In the calculation of the fitness value, the MOHEA needs $mN(N-1)/2$ comparisons to calculate PDDR-FF value, so the computational complexity of MOHEA is $O(mN \log_2 N)$. The computational complexity of NSGA-II is the same as SPEA2. The value is $O(mN^2)$. In summary based on the quantitative analysis of the complexity, the MOHEA is clarified faster than NSGA-II and SPEA2, and the experimental result are shown in Table 1.

5 Conclusions

In this study, a multi-objective hybrid evolutionary algorithm (MOHEA) based on mixed sampling strategy was proposed to solve the multi-objective optimization problem while improving the convergence and distribution performances, and reducing the computational time at the same time. The VEGA sampling strategy prefers the edge region of the Pareto front, but deviates from the search of the central region. The PDDR-FF sampling strategy makes up the shortcoming of VEGA. The mixed sample strategy could converge the multiply regions of Pareto front to improve the efficacy. Since no computation requirement for distance, the MOHEA could also improve the efficiency. Numerical comparisons indicated that the proposed MOHEA outperforms the classical MOEAs such as SPEA 2 and NSGA ii on convergence and distribution performance as well as computational time too.

Acknowledgements. This research work is supported by the National Natural Science Foundation of China (U1304609), Foundation for Science & Technology Research Project of Henan Province (162102210044, 152102210068, 152102110076), Program for Innovative Research Team (in Science and Technology) in University of Henan Province (17IRTSTHN011), Program for Key Project of Science and Technology in University of Education Department of Henan Province (17A520030), Fundamental Research Funds for the Henan Provincial Colleges and Universities (2014YWQQ12, 2015XTCX03, 2015XTCX04), Research Funds for Key Laboratory of Grain Information Processing and Control (Henan University of Technology) (KFJJ-2015-106), Ministry of Education and the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS): No. 15K00357.

References

1. Deb K, Beyer HG (2001) Self-adaptive genetic algorithms with simulated binary crossover. Evol Comput 9(2):197–221
2. Deb K, Goyal M (1999) A combined genetic adaptive search (geneas) for engineering design. Comput Sci Inform 26:30–45
3. Deb K, Pratap A et al (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6(2):182–197
4. Gaspar-Cunha A, Covas JA (2001) Robustness in multi-objective optimization using evolutionary algorithms. John Wiley & Sons, Inc
5. Li H, Zhang Q (2009) Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II. IEEE Trans Evol Comput 13(2):284–302
6. Non-Member WZ, Member SFS (2012) Multiobjective process planning and scheduling using improved vector evaluated genetic algorithm with archive. IEEJ Trans Electr Electron Eng 7(3):258–267
7. Schaffer JD (1985) Multiple objective optimization with vector evaluated genetic algorithms. In: International Conference on Genetic Algorithms, pp 93–100
8. Schott JR (1995) Fault tolerant design using single and multicriteria genetic algorithm optimization. Cell Immunol 37(1):1–13
9. Tukey JW (1978) Variations of box plots. Am Stat 32(1):12–16

10. Yao X, Liu Y, Lin G (1999) Evolutionary programming made faster. IEEE Trans Evol Comput 3(2):82–102
11. Yu X, Gen M (2010) Introduction to Evolutionary Algorithms. Springer, London
12. Zitzler E, Thiele L (1999) Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. IEEE Trans Evol Comput 3(4):257–271
13. Zitzler E, Laumanns M, Thiele L (2001) Spea2: Improving the strength pareto evolutionary algorithm. vol 3242, pp 95–100

Literature Mining Based Hydrogen Fuel Cell Research

Luoji Li[1], Ying Zhang[2](✉), and Qiulin Li[3]

[1] Shenyang Sport University, Shenyang 110102, People's Republic of China
[2] Environment School, Liaoning University, Shenyang 110036, People's Republic of China
yingying0809@126.com
[3] Business School, Sichuan University, Chengdu 610065, People's Republic of China

Abstract. In this paper, we reviewed the current technology for the hydrogen fuel cells and got the light spots when 1271 papers related to hydrogen fuel cells were analyzed, which were focused on the hydrogen fuel cell technologies, hydrogen fuel cell application and solar hydrogen fuel cell. This paper had outlined these several types of the specific technology and confirmed that hydrogen fuel cells was an inevitable choice when faced to serious environmental pollution and energy structure's imbalance. And the solar-hydrogen fuel cell system that involved the hydrogen production, hydrogen storage and hydrogen application is of significant importance addressing the problem of the production of renewable energy, power generation and transportation pollution.

Keywords: Hydrogen fuel cell · Solar-hydrogen · FC vehicles · Zero-pollution

1 Introduction

Many environmental issues have been caused by or relate to the production, transformation and use of fossil energies, for example, acid rain, stratospheric ozone depletion and global climate change [5]. Energy is the material basis for the country's social development and progress of science and technology, but due to the constant use of the three large fossil energies, energy conservation and environmental problem became more and more serious. As a result of the badly exceeds bid of PM2.5, human beings got a warning of continuous haze weather. And the air pollution problem can be settled down when it is mainly closely related to fossil fuel combustion and transport industries. The increasing trend in world's energy need is expected to continue in the future. As a result, a growth in energy generation capacity will be needed [1]. Cars are important transport means in people's lives. However, the automobile exhaust is the important cause of increasingly serious environmental pollution. Therefore, we need to find an alternative fuel for the traditional fossil fuels.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_9

Hydrogen has high energy density, and the energy released is enough to make the vehicle's engine running. Hydrogen energy systems appear to be the one of the most effective solutions and can play a significant role in providing better environment and sustainability [4,14]. Hydrogen fuel cell vehicle itself works with no noise, no vibration, no loss, and long service life. The electrode is only working as the place of chemical reaction and conductive channel and it does not participate in chemical reactions. The fuel of hydrogen fuel cell is hydrogen and oxygen, and the product is clean water, with the prefect results of no CO and CO_2, and no sulfur and particulate exhaust. We could have a clear look at the advantages when compared to the thermal power as Table 1 shows. Therefore, hydrogen fuel cell vehicle can run with zero emissions, zero pollution in real sense and hydrogen is regarded as a perfect energy for vehicle. Fuel cell technology with the hydrogen energy is the core technology of the auto industry in the 21st century, and it takes revolutionary significance for the car industry.

Table 1. Comparison of atmospheric pollution between fuel cells and thermal power

Pollution components	Gas thermal power	Heavy oil thermal power	Coal thermal power	Fuel cells
SO_2	2.5–230	4550	8200	0–0.12
NO_2	1800	3200	3200	63–107
Hydrocarbon	20–1270	135–5000	30–10000	14–102
Dust	0–90	45–320	365–680	0–0.14

(unit: kg·10^{-6}(KWh)$^{-1}$)

In the face of oil price's rising, increasingly severe atmospheric environment, increasingly strict emission regulations and the pressing needs of light energy, fuel cell is the power generation device which can turn the chemical energy stored in the fuel and oxidant directly into electrical energy. It is considered to be the best choice as clean and efficient power generation technology in the 21st century because of its energy conversion is not limited by Carnot cycle [8], energy conversion efficiency can reach 90%, and the actual service efficiency is 2–3 times of internal combustion engine. And hydrogen fuel cell shows the unique superiority in such aspects as in the comprehensive energy efficiency, environment friendly and high reliability, so the hydrogen fuel cell will have broad market prospects as specific tools for the transition from fossil fuel based era to hydrogen economy-oriented society.

2 Literature Mining

Hydrogen fuel cell is a cutting-edge research problem in current society, but researches about this issue has certain basis and achievements. Based on these researches, literature mining has been shown to be a powerful method for elucidating major trends across time in published scientific literature so that topic maps can be built [2]. Garfield believed that the academic literature citation indexing was crucial when researching similar topic areas [7]. A citation index is a synthesized result based on journal articles, keywords, publication dates and abstracts and is able to separate and highlight the various influences in a specific field, allowing for the research with the greatest impact to be easily identified.

2.1 Data Analysis

Literature mining from published scientific literature allows for the discovery of key areas and trends [2]. CiteSpace is one of the latest development of a generic approaches to detecting and visualizing emerging trends and transient patterns in scientific literature. The work makes substantial theoretical and methodological contributions to progressive knowledge domain visualization. A specialty is conceptualized and visualized as a time-variant duality between two fundamental concepts in information science: research fronts and intellectual bases [3]. 1271 papers had been retrieved as "TI = HFC" been limited, and all the data was imported into the CiteSpace with a data process. Finally, 65 categories were formed after cluster, as Fig. 1 shows.

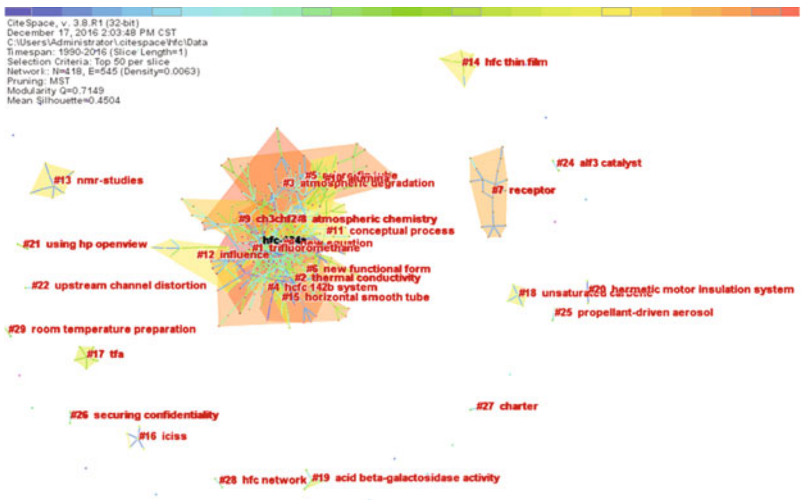


Fig. 1. Clusters of the hydrogen fuel cells

2.2 Results

Keywords are important index terms used to provide access to articles that have been published or presented in journals and conference databases and are essential for clearly identifying arange within a specified boundary. As Fig. 2 illustrates, keywords about hydrogen fuel cells have been shown from 1960 to 2016, which could be divide into three phases as timezone developed. And the first stage could be described as hydrogen fuel cell technologies in Sect. 3, the second could be extracted as hydrogen fuel cell application in Sect. 4, and the latest stage could be know as solar hydrogen fuel cell in Sect. 5, which is on behalf of the most frontier and hot issue on hydrogen fuel cells.

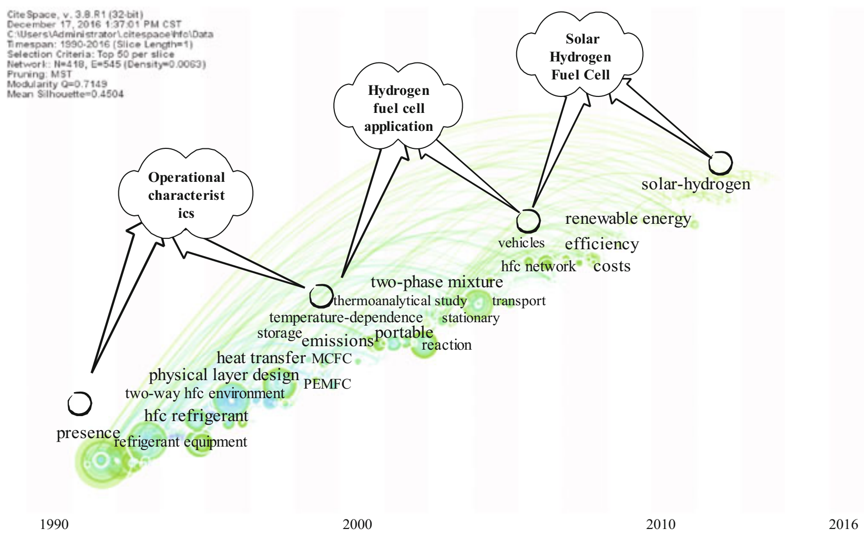


Fig. 2. Timezone of the hydrogen fuel cells

3 Hydrogen Fuel Cell Technologies

Hydrogen production has laid the basis for the development of fuel cells (FC) and the diffusion of associated technologies. The fuel cell technology mechanisms were discovered in 1838 by the German-Swiss chemist Christian Friedrich Schonbein and the British lawyer and natural scientist Sir William Grover [17]. The FC technologies have the potential to be the key technologies that contribute to the transition to a low carbon energy economy. The fuel cell and its application has become the most attractive and promising technology in the hydrogen energy utilization field. Because of its' promising clean and efficient power generation, as oil and the internal combustion engine displaced coal and the steam engine, FCs will make the transition in a similar way in the future.

3.1 Operational Characteristics

A fuel cell is similar to conventional battery but different, being composed of an anode, cathode and electrolyte. These components generate electricity through an electrochemical reaction in which the external oxygen supply source (usually air) and a hydrogen-rich fuel combine to form water as Eq. (2) show. The difference with a conventional battery is that no combustion occurs in FCs, as the energy to run the FCs is released electro-catalytically. This electrochemical reaction at the two electrodes is shown from Eqs. (1) to (2) respectively.

$$\text{Atanode:} \, H_2 \rightarrow 2H^{+}{+}2e^{-}., \tag{1}$$

$$\text{Atcathode:} \, 2H^{+}{+}2e^{-}{+}O_2 \rightarrow \text{H2O}. \tag{2}$$

As the schematic diagram in Fig. 3 shows, the fuel (hydrogen) and oxidizer (oxygen) dissociate to evolve electrons and ions, the ions separating from the pole move towards another under electrolyte conditions while the dissociated electrons are guided to an external electric circuit. The fuel (hydrogen) is supplied for the cathode, and the oxidizer for the anode. The hydrogen dissociates the H^+ and e^-, and the H^+ blends into the electrolyte while the e^-, tends towards the positive pole through an external circuit. The H^+ and oxygen in the air react at the positive pole to form water when absorbing e^-,. As long as the fuel and oxygen are supplied, there is a constant power output in the FCs.

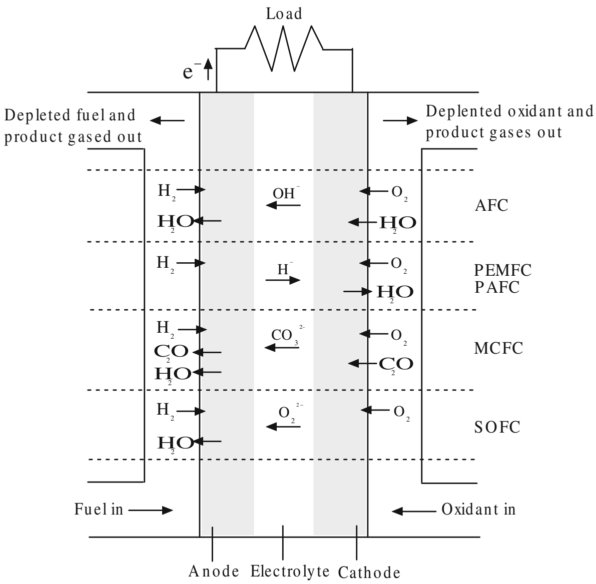


Fig. 3. The schematic diagram of FCs and different types FCs' operation principle

3.2 Outlook for Fuel Cell Technologies

FCs are classified by the kinds of electrolytes they use, which determines the kind of chemical reactions that take place in the cell, the kind of catalysts required, the temperature range in which the cell operates, the fuel required, and other factors. In turn, these chemical reaction condition characteristics affect the applications for which these cells are most suitable. There are six major types of fuel cells presently under development, each with its own advantages, limitations, and potential applications, as shown in Table 2. Respectively, they are proton exchange membrane fuel cells (PEMFC), alkaline fuel cells (AFC), phosphoric acid fuel cells (PAFC), molten carbonate fuel cells (MCFC), solid oxide fuel cells (SOFC) and direct methanol fuel cells (DMFC). The operating characteristics of these systems are highlighted in Table 2.

Table 2. Operating and applicable properties of FCs technologies

Styles	PAFC	MCFC	SOFC	PEMFC	AFC
Fuel	Natural gas, methane hydrogen peroxide, air	Natural gas methane gas, air, coal gas, hydrogen peroxide	Natural gas methane gas, coal gas, air, hydrogen peroxide,	Hydrogen, oxygen	Hydrogen, oxygen
Electrolyte	phosphate	Alkali metal carbonate fused mixture	Oxygen ion conductive ceramics	Proton exchange membrane	Potassium hydroxide solution
Operating temperature	160–220	620–660	800–1000	80	90
Electrochemical efficiency	55%	65%	60–65%	40–60%	60–70%
Power output	200KW	2MW–10MW	100KW	1KW	300W–5KW

AFCs use an aqueous solution of potassium hydroxide (KOH) as an electrolyte with a concentration ranging from 35% to 45%, a fuel of pure hydrogen, and typically operate at a low temperature of around 70° C. It is a one of the most developed technologies so far and has been used by NASA spacecraft. Because of the low operating temperature, instead of a platinum catalyst in the system, a variety of non-precious metals such as the most commonly used catalyst of nickel can be employed to speed up the reactions taking place at the anode and cathode in AFC units. Owing to the rate at which the chemical reactions take place, AFCs offer relatively high fuel to electricity conversion efficiencies, as high as 60% in some applications. The major advantage of AFCs is the quick start, but the disadvantage is that they are very sensitive to CO_2 which causes a slower reaction time and consumes the alkaline in the electrolyte, thereby limiting its commercial applications.

PEMFCs are considered to be a promising technology for clean and efficient power generation in the 21st century and are known for their high-power density delivery and short start up time compared with other fuel cells, characteristics which guarantee that the PEMFCs technology is competitive for transportation and commercial applications such as stationary, and portable power generation [13]. PEMFCs use a solid polymer electrolyte (Teflon-like membrane) which is an excellent conductor of protons and an insulator for electrons to exchange ions between two porous electrodes, and operate under a temperature of 100° C. The operating schematics are shown in Fig. 3. Therefore, the membrane is the key point in the PEMFCs technology and current research on PEMFCs also focuses on the development of a proton exchange membrane with high proton conductivity, low electronic conductivity, good thermal stability, and low cost.

Figure 3 shows that the working principles of the PAFCs and PEMFCs are similar even though the structures are different. PAFCs use liquid phosphoric acid as an electrolyte which is contained in a Teflon-bonded silicon carbide matrix, and works at about 150 to 200° C. Different from PEMFCs and AFCs, PAFCs are more tolerant and insensitive to any impurities in the hydrogen fuel. The advantages of PAFCs are that the metal catalysts required are much less than that of the AFCs and the reducing agent purity requirements are significantly lower allowing for up to 5% carbon monoxide. PAFCs are viewed as the "first generation" of modern FCs. According to the state development, these are the first mature cell types to be used commercially. The 100200 and 500 kW size plants are typically available for stationary power generation. A 1.3 MW system has already been tested in Milan and PAFCs have been installed at 70 sites in Europe, USA and Japan [16].

MCFCs are high-temperature (650° C) fuel cells which commonly use a molten carbonate salt, such as lithium carbonate, potassium carbonate or sodium carbonate, suspended in a porous ceramic matrix as the electrolyte. To date, this type has been developed in natural gas and coal based power plants for industrial and military use. The SOFCs' electrolyte is a hard, non-porous ceramic compound such as yttria stabilized zirconia which has a strong conductivity and operates at a much higher temperature of 800–1000° C [12]. SOFCs are considered to be around 50%–60% efficient at converting fuel to electricity and have rapidly increased in popularity in stationary applications in recent years.

Different from the five types fuel cells which fueled by hydrogen, DMFCs are powered by pure methanol or alcohol which can be generated hydrogen via reforming reaction as presented in Sect. 3. Direct methanol fuel cell technology is relatively new compared with that of fuel cells powered by pure hydrogen, and DMFCs research and development is roughly 3–4 years behind that for other fuel cell types [10]. The working principle and reaction equations are shown in Eqs. (3)–(4).

$$\text{Atanode: } CH_3OH + H_2O \rightarrow CO_2 + 6e^- + 6H^+, \tag{3}$$

$$\text{Atcathode: } \frac{3}{2}O_2 + 6e^- + 6H^+ \rightarrow 3H_2O. \tag{4}$$

4 Hydrogen Fuel Cell Application

Generally, today, the application of fuel cell technologies can be categorized in three directions. Respectively, there are portable fuel cells designed to be moved, including auxiliary power units (APU), stationary power fuel cells designed to provide power to a fixed location, and transport fuel cells providing either primary propulsion or range-extending capacity for vehicles as shown in Table 3.

Table 3. A brief summary of hydrogen energy application technical directions [11]

Application	Portable	Stationary	Transport
Definition	Units that are built into, or charge up, products that are designed ti be moved, including auxiliary power units (APU)	Units that provide electricity sometimes, but are not designed to be moved	Units that provide propulsive power or ranger extension to vehicle
Typical power range	5W–20KW	0.5KW–400KW	1KW–100KW
Typical technology	PEMFE, DMFC	MCFC, PAFC, PEMFC, SOFC	PEMFC, DEMFC
Examples	Non-motive APU (campervants, boats, lighting); military applications; portable products; small personal electronics	Large stationary combined heat and power; small stationary micro-CHP; uninterruptible power supplies(UPS)	Materials handing vehicles; fuel cell electric vehicles (FCEV); Trucks and buses

4.1 Portable Technological Direction

Using portable fuel cells can solve the low energy power capability and long charging time problems of current battery technology. Consequently, the global production of portable fuel cells has grown continuously on the basis of FC technology. In the 21st century, commercial portable units are being adopted rapidly in such areas as educational fuel cells and toys, auxiliary power units (APU), and consumer electronics [11]. With the many specific industrial niches that can benefit from fuel cells, the technology's use in APU has been an area of interest for fuel cell manufacturers and industry players alike. DMFCs are a relatively expensive cell technology and is most cost-effective in small, low-power stack configurations under 100 W. There are a number of products emerging for applications that require a higher APU output. The Danish company Serenergy uses methanol as a fuel, but processes it ahead of a 350 W high-temperature PEMFC (HT PEMFC), which it is marketing fora variety of applications.

The long-awaited launch of a number of fuel cell consumer electronics chargers in 2011 and 2012 has marked a turning point in this sector. For many years it proved too difficult to miniaturize fuel cell technology to a level where it could realistically compete with external battery supplies. Continued development is expected to result in units small enough to be integrated inside consumer devices and large multinational companies, such as Apple and Research In Motion, are patenting products in this field.

4.2 Stationary Technological Direction

Stationary units that provide electricity/heat are not designed to be moved. They are used for backup power, in remote locations, as stand-alone power plants for towns and cities, for distributed generation in buildings, and for co-generation [9]. Compared with traditional power stations, stationary fuel cells use is inevitable because of today's subtle change in energy structures and the breakthroughs in fuel cell technology.

In recent years, both small and large systems in residential, industrial and prime power applications have experienced a fast growth such as in the highly-advanced Japanese Ene-Farm residential micro-CHP plan and the emergence of similar schemes in Europe. Present fuel cell technology allows for a continued cost reduction and optimization of the technology. At the same time, the commercialization of backup power systems for the telecommunications industry has accelerated as the communication field expands with a significant interest emerging in China and elsewhere for both independent and grid-connected systems. The introduction of methanol-fuelled telecommunication backup locations, where delivering hydrogen is difficult, to take advantage of fuel cell technology. Large stationary fuel cell installations have gained increasing interest from multinational companies who are looking for ways to reduce their carbon footprint.

4.3 Transport Technological Direction

The global automotive industry sells in excess of 80 million vehicles per year, and for fuel cells a portion of this market remains the long-term prize in this sector. FCs are capable of replacing ICEs (internal-combustion engines) as they have the potential to achieve higher efficiency and lower GHG emissions. A typical application in vehicles varies from passenger cars, utility vehicles, to buses, with power ranges from 20 kW to 250 kW [15].

The production of regular automobiles increased steadily in the early 2000 s. In the past few years, the fuel cell light-weight vehicle market has been led by Honda, General Motors, and others. Honda has started shipping its FCX Clarity to Southern and Northern California. In 2007 General Motors, through its "Project Driveway" program, delivered over 100 units of its Chevrolet Equinox fuel-cell vehicles to California, Washington DC, and New York. Other fuel cell application technologies are in marine and unmanned aerial vehicles (UAVs), trucks, trains, and bicycles. Siemens has developed a fuel-cell-based air-independent propulsion system intended for use in submarines, such as the U212/214 type hybrid powered submarine manufactured by the German Howaldtswerke Deutsche Werft GmbH (HDW).

5 Solar Hydrogen Fuel Cell

Solar energy is a ubiquitous and rich source, also a kind of renewable energy sources with no environmental pollution. China has abundant solar energy resource, annual total solar radiation area is greater than 1050 kWh/m^2, which is more than 96% of the total land. There are several methods for producing hydrogen from solar energy. Currently, the most widely used solar hydrogen production method is to obtain hydrogen by electrolyzing the water at low temperature. Solar-hydrogen fuel cell power generation system was new-type energy which was based on the solar power for water decomposition, and it could generate electricity by electrochemical reaction using the hydrogen and oxygen produced. Countries are committed to carry on the development and application of solar-hydrogen fuel cell because of rich energy source, high energy conversion efficiency, and no pollution in the whole transformation process. In the future energy system, solar energy will become the main primary energy to replace the current energy such as the coal, oil and gas, and hydrogen fuel cell will become the clean energy to replace petrol, diesel and chemical battery.

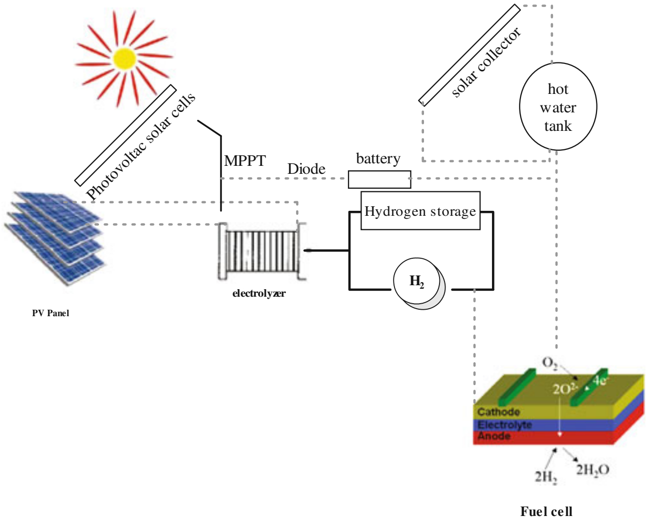


Fig. 4. Scheme of the solar-hydrogen system

As a reversible fuel cell, it gives you the freedom to invent your own clean energy applications using fuel cells and renewable hydrogen formed using sun and water. Figure 4 shows a schematic diagram of hydrogen production system that uses photovoltaic technology. The electrolysis of distilled water using electricity produced by the PV panel takes place in electrolyser unit and produces hydrogen and oxygen as the end product [6]. Through the study of solar-hydrogen fuel cell system's characteristics, we known that hydrogen fuel cell power generation efficiency could be as high as 36.4%, and solar cell power generation efficiency generally is in 15% to 18% range. Therefore, the hydrogen fuel cells not only greatly improve the efficiency of the power generation, also produce only water in the whole power generation process, which implement the clean, pollution-free and efficient, and will become a leader in the field of new energy in the 21st century.

6 Conclusion

Through the researches on the existing hydrogen fuel cell technologies and the application technologies, we learned that the real value of the hydrogen fuel cells is does not poor. Especially this particular and reversible solar-hydrogen fuel cell system is bright star prior to other styles with a steady stream of resources and the cycle of zero pollution. The direct hydrogen fuel cell vehicle is preferred, since it would be less complex, have better fuel economy, lower greenhouse gas emissions, greater oil import reductions and would lead to a sustainable transportation system once renewable energy was used to produce hydrogen. All countries need to increase comprehensive and multi-functional analysis and research on hydrogen fuel cell's characteristics if we want to keep in active status in the field of energy development in the 21st century.

Acknowledgements. This research is supported by the Liaoning economic and social development project (No. 2017slktjd-063); The authors are also grateful to the Liaoning doctoral research start-up funding (No. 201501131).

References

1. Acar C, Dincer I (2014) Comparative assessment of hydrogen production methods from renewable and non-renewable sourcesn. Int J Hydrogen Energy 39(1):1–12
2. Bruijn BD, Martin J (2002) Getting to the (c) ore of knowledge: mining biomedical literature. Int J Med Inform 67(1):7–18
3. Chen C (2006) CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. J Am Soc Inform Sci Technol 57(3):359–377
4. Dincer I (2007) Environmental and sustainability aspects ofhydrogen and fuel cell systems. Int J Hydrogen Energy 31(1):29–55
5. Dincer I (2012) Green methods for hydrogen production. Int J Hydrogen Energy 37(2):1954–1971
6. Joshi AS, Dincer I, Reddy BV (2010) Exergetic assessment of solar hydrogen production methods. Int J Hydrogen Energy 35(10):4901–4908
7. Leydesdorff L (2015) Bibliometrics/citation networks. arXiv preprint arXiv 1502.06378
8. Li J, Chen L et al (2015) Finite time optimizations of a newton's law carnot cycle. Int J Energy Environ 5(6):517

9. Miller EL, Randolph K et al (2016) Innovative approaches to addressing the fundamental materials challenges in hydrogen and fuel cell technologies. MRS Advances 1:1–13
10. Murray EP, Tsai T, Barnett SA (1999) A direct-methane fuel cell with a ceria-based anode. Nature 400(6745):649–651
11. Narayanan S, Valdez T (2002) Portable direct methanol fuel cell systems. Handbook of fuel cells
12. Park SD, Vohs JM, Gorte RJ (2000) Direct oxidation of hydrocarbons in a solid-oxide fuel cell. Nature 404(6775):265–267
13. Stumper J, Stone C (2008) Recent advances in fuel cell technology at ballard. J Power Sources 176(2):468–476
14. Turner J, Sverdrup G et al (2008) Renewable hydrogen production. Int J Energy Res 32(5):379–407
15. Wilberforce T, Alaswad A et al (2016) Advances in stationary and portable fuel cell applications. Int J Hydrogen Energy 41(37):16,509–16,522
16. Winter CJ (2009) Hydrogen energy-abundant, efficient, clean: a debate over the energy-system-of- change. Int J Hydrogen Energy 34(14):1–52
17. Young S (2001) Tomorrow's energy: hydrogen, fuel cells and the prospects for a cleaner planet. Nature 414(6863):487–488

GM(1.1) Model Based Leshan's Industries Shift-Share Analysis and Prediction

Hongxing Lan[1,2], Caiyang Xu[1], and Zhiyi Meng[1,2](✉)

[1] Business College, Sichuan Agricultural University, Dujiangyan 611800, People's Republic of China
zhiyimengscu@sina.com

[2] Southwest Poverty Reduction Center, Sichuan Agricultural University, Chengdu 610065, People's Republic of China

Abstract. The technique of shift-share method [6] and GM(1.1) model [4] are applied for a dynamic assessment and predictive study of industries between the year of 2010 and 2015 in Leshan city, choosing Sichuan province as a background. With the actuality, some advice was put forward on how to adjust and optimize the development of industries of Leshan. The survey shows that the secondary industry is the key factor to the whole economic development of Leshan City. Leshan's tertiary industries had a weak competitiveness and will develop more slowly, department structure contributed a little to the rapid growth of economic total; Tourist industry and trade are lack of competitiveness, while the transportation industry industries ran on a small scale and contribute little to the growth of economic total. Hence, it's important and necessary to formulate the industry development strategy based on objective scientific analysis of Leshan's industries.

Keywords: GM(1.1) model · Shift-share · Leshan city · Industrial policy

1 Introduction

Shift-Share Method (SSM) was put forward by America scholars Dunn, Perloff, Lampard and Muth in the 60s, and it was summarized as the form which is widely acknowledged now by Dunn in the early 80s [2]. Shift-Share Method is widely applied in the assessment of regional economy and industrial analysis with high comprehensiveness and dynamic which is used mainly in analyzing regional economy efficiency. Due to the particularity of object, every practical application has its own referring value.

SSM was introduced to China in the 80s. Nowdays lots of domestic scholars study in inter-region with SSM widespread and some surveys already found that it is evident and suitable for the area of the industrial development assessment [1,5,9]. However, the utilizing of SSM by existing research is based on the analysis

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_10

of economic development, depending on the experience to predict the growing trend of the future whose forecast of the trend still has a high level of subjectivity. As a result, this text applies GM(1.1) model to predict the chief data of industrial development. To employ this predicative study to estimate the development of regional industry in the future on the basis of dynamically assessing industrial development with SSM.

Following a longitudinal study, we can explore the rule of industry development to provide the adjustment of regional industry organization and industry policy formulation with more scientific decision source.

Leshan city is a significant connection point and port of the economic zones of the southern Sichuan and Chengdu areas, also the west of Panzhihua which plays a vital role in economic development structure. On one hand, Leshan city has an outstanding advantage in agriculture, industry, tourism, transportation and business. On the other hand, the problems of inadequate use of ascendant resource still exist. From the view of whole Sichuan province, the overall economic development of Leshan city remains an upper medium level; in the perspective of industrial department, the superior industries are not obvious enough which cannot be judged by experience to find out the departments with advantages and growth potential. Energy-extensive consumption, high-emission and high-pollution still have an influence on the economic development of Leshan city. In this situation, if efforts are made in pursuit of long-term development, it will solve the problem of increasing load of resource and environment. Therefore, getting a clear understanding of industrial development trend will offer reliable evidence to the transformation of low-carbon industry. As a result, it is of great realistic significance to apply the technique of SSM, dynamic assessment and predictive study to industrial development of Leshan city.

2 Research Methods and Data Sources

2.1 Shift-Share Method and GM(1.1) Model

SSM regards regional economy as a dynamic process and views its locating region or whole country as a reference system. Then this method divides the variation of the regions' own economic aggregate into three subprojects, which are the national growth effect, the industry mix effect and the competitive effect to approve the reason of this region's economy development and recession. At the same time, it also can evaluate the strength of regional economic structure's competitiveness. $N_{ij} = b_{ij} \times R_j$ is called as national growth effect which means Department J's total quantity is distributed pro rata. The variation of Department J's scale in area "i" is the variation caused by regional industrial department for standardization in average growth rate of locating area or whole country. $P_{ij} = (b_{ij,0} - b_{ij}) \times R_j$ is marked as the industry mix effect (industry structure effect) which reflects the influence and contribution from the structure of department to the increase: the rate is greater and that contribution is greater. $D_{ij} = b_{ij,0} \times (r_{ij} - R_j)$ is named as the competitive effect (the local share effect) which shows the relative competitiveness of Department J, so that a greater rate means a greater effect to economic increase made by Department J's competitiveness, including $r_{ij} = \frac{b_{ij,t} - b_{ij,0}}{b_{ij,0}}$, $R_j = \frac{B_{j,t} - B_{j,0}}{B_{j,0}}$, $b_{ij} = \frac{b_{ij,0} \times B_{j,0}}{B_0}$. Introduce $K_{j,0} = b_{ij,0}/B_{j,0}$, $K_{j,t} = b_{ij,t}/B_{j,t}$ which means the relative department's share of the whole area or country at the beginning and at last. The relative increasing rate of area i, compared with its locating area or country, is $L = W \times U$,

$$W = \frac{\sum_{j=1}^{n} K_{j,0} \times B_{j,t}}{\sum_{j=1}^{n} K_{j,0} \times B_{j,0}} : \frac{\sum_{j=1}^{n} B_{j,t}}{\sum_{j=i}^{n} B_{j,0}}, \quad U = \frac{\sum_{j=1}^{n} K_{j,t} \times B_{j,t}}{\sum_{j=1}^{n} K_{j,0} \times B_{j,t}},$$

where W and U separately represent result the effect index and effect of regional competition index. If G_i is a little bit big while L is greater than 1, the regional increase is faster than another which is of locating area and whole country. If P_i is a little bit large while W is greater than 1, it proves the industry department with high-speed growth has major share. In this situation, the regional overall economic structure is a little bit good and the structure makes a great contribution to economic growth. If D_i is great and U is greater than 1, the increasing trend of every industry department is obvious with a strong competitiveness.

GM(1.1) model has part of the information known and partial information unknown. It can accurately describe the state and behavior of the social economic system. On one hand, this method can help avoid the Achilles' heel of the lack of relevant data. On the other hand, it also can avoid the subjective assume due to personal preference, experience, knowledge and macro policy which can better grasp the self-evolution law of the system.

GM(1.1) was modeled as follows:

(1) Set the original sequence $X^{(0)}$ and figure up the original sequence according to $x^{(1)}(i) = \sum_{m=1}^{i} x^{(0)}(m), i = 1, 2, \cdots, n$ to create a new sequence $X^{(1)}$;
(2) Create a winterization differential equation [12] GM(1.1) as predictive model to sequence $X^{(1)}$, that is $\frac{dX^{(l)}}{t} + aX^{(l)} = b$, a, b are undetermined parameter, separately as evolution parameter and gray action;
(3) Suppose $\hat{a}$ as undetermined parameter vector, with the help of least square method, to figure out $\hat{a}$ and get $\hat{a} = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T X$;

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}[x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{bmatrix}, \quad X = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

(4) Figure out GM(1.1) and get $\hat{X}^{(1)}(i+1) = (x^{(0)}(1) - \frac{b}{a})e^{-ai} + \frac{b}{a}$;
(5) According to Inverse Accumulated Generating Operation to calculate $x^{(0)}(i+1)$; Make a counter current to the predictive value $\hat{X}^{(0)}(i+1) = \hat{X}^{(1)}(i+1) - \hat{X}^{(1)}(i)$;
(6) Make an error checking to prediction model, including residual test, relational coefficient test and posterior-variance-test [11].

2.2 Definition of Indicator

In recent years, Leshan city's GDP has been maintaining a double-digit growth in 10 years whose comprehensive strength significantly strengthened. In 2009, the domestic GDP is 61.902 billion yuan–15.3% more than the number of the previous year. Three industries separately stimulate GDP growth by 0.5%, 11.1%, 3.7% whose contribution rate of economic growth separately are 3.1%, 72.6%, 24.3%. It is necessary to analyze the three industries and they are marked as GDP1, GDP2, GDP3 in the sections below. During The Eleventh Five-Year-Plan, by building "the transportation project with one hundred billion yuan", developing the economy of Culture & Tourisim, business, mountain area and port-surrounding, it promoted the position as a south-west hub and accelerated the development of economy and society. Leshan city is a vital city of industry and tourism of Sichuan province. At the same time, it belongs to the core area in Chengdu economic zone as an important city of business, so that its tourism, transportation industry and business are regarded as leading industries for economic development. Therefore, in the current study, based on the analysis of Leshan city's three industries, for further study, choose tourism, transportation and storage postal service, business industry from thrice industry to analyze. In consideration of data's access, the current study chooses Total Retail Sales of Consumer Goods as an alternative to reflect the prosperity degree of the business industry. Tourism is marked as GDP31 while transportation storage and transport the postal service is marked as GDP31. SXT shows Total Retail Sales of Consumer Goods.

2.3 Data Sources

The data come from *Leshan City Statistical Yearbook* (1999–2009) [3] and the documents from relevant departments; *Sichuan Province Statistical Yearbook* (1999–2009) [8]. The data are calculated in current prices.

3 The Analysis and Prediction of Industry Development Share-Shift

(1) The Prediction of Industry Development
According to GM(1.1) model, every index will be predicted and calculated to get Index Prediction as follow Table 1:

Get relative error sequence of every index $\Delta_i = \frac{\varepsilon(i)}{x^{(0)}(i)}$ and make a residual test to model. If average relative error is not more than 5%, its predicted value meets the precision requirement. Specific contents are as follows Table 2.

As shown in Table 2, the average relative error of the tourism industry in Sichuan province, the tourism, transportation storage and transport the postal service of Leshan are more than 5% which cannot pass the residual test. On the basis of the analysis of the relative error every year, the above three indicators in 2002–2004 relative error is a little bit obvious while the error of 2003 is

Table 1. 2010–2015 The prediction of industrial data of Leshan city (L) and Sichuan province (S)

	2010	2011	2012	2013	2014	2015
S GDP1	2730.12	3073.91	3461	3896.83	4387.55	4940.06
S GDP2	8298.7	10117.14	12334.04	15036.72	18331.62	22348.51
S GDP3	5941.44	6896.16	8004.29	9290.49	10783.36	12516.12
L GDP1	5941.386	6896.162	8004.37	9290.67	10783.67	12516.6
L GDP2	462.0822	567.4739	696.9	855.85	1051.06	1290.78
L GDP3	194	226.4765	264.38	308.63	360.28	420.58
L GDP31	1805.09	2157.87	2579.59	3083.73	3686.39	4406.84
S GDP32	595.91	635.097	676.85	721.36	768.79	819.34
S SXT	6704.55	7893.34	9292.92	10940.66	12880.56	15164.42
L GDP31	139.28	164.66	194.66	230.16	272.04	321.6
L GDP32	20.95	24.69	29.1	34.29	40.4	47.61
L SXT	269.35	316.56	372.04	437.24	513.88	603.94

Table 2. GM(1,1)Residual test

Index	Average relative error	Index	Average relative error
S GDP1	5.00%	S GDP31	11.22%
S GDP2	1.68%	S GDP32	3.36%
S GDP3	1.28%	S SXT	2.30%
L GDP1	5.71%	L GDP31	11.13%
L GDP2	4.57%	L GDP32	10.09%
L GDP3	2.87%	L SXT	1.63%

Relative error	Average	2002	2003	2004	2005	2006	2007	2008	2009
S GDP31	11.22%	13.84%	23.00%	9.24%	2.51%	9.77%	13.20%	15.63%	2.55%
L GDP31	11.13%	14.82%	21.18%	9.41%	1.17%	10.99%	16.50%	9.02%	5.97%
L GDP32	10.09%	14.75%	11.35%	15.50%	8.89%	10.20%	6.06%	5.81%	8.16%

the largest–the prediction relative error between Leshan and Sichuan province in tourism is more than 20% due to the outbreak of SARS which makes a profound impact [7]. Above conclusion is also mentioned in previous studies and GM(1.1) model still cannot eliminate the phenomenon effectively. Therefore, on the basis of acknowledging prediction error and considering the reality, the three sets of predictive data meets the precision requirement.

Calculate the correlation coefficient, correlation degree (while $\rho = 0.5$, $\zeta > 0.6$ is satisfying), variance ratio (if $C < 0.35$, and $P \geq 0.95$, the forecasting precision is regarded as superior). The result shows as follow Table 3: the correlation degree of every index is greater than 0.6; the variance ratio of every index is less than 0.35; every small error possibility is 1. According to the inspection standard, this prediction accuracy is higher and predicted values have high credibility.

Table 3. GM(1,1) Correlation degree and Posterior-variance-test

Project	Correlation degree	C	P	Project	Correlation degree	C	P
S GDP1	0.648	0.148	1	S GDP31	0.663	0.142	1
S GDP2	0.762	0.035	1	S GDP32	0.615	0.174	1
S GDP3	0.727	0.019	1	S SXT	0.671	0.026	1
L GDP1	0.623	0.113	1	L GDP31	0.627	0.152	1
L GDP2	0.762	0.087	1	L GDP32	0.875	0.066	1
L GDP3	0.663	0.056	1	L SXT	0.689	0.025	1

(2)The Analysis of Industry Development Share-Shift
According to the SSM model and using Sichuan province as a reference area, analyze and predict the share-shift of Leshan industrial data from 2002 to 2015 and the results are showed as follow Tables 4, 5 and 6.

Table 4. 2002–2009 The SSM analysis of three industries of Leshan city

Index	Leshan	2002	2003	2004	2005	2006	2007	2008	2009
P_{ij}	GDP1	2.07	2.641	9.606	3.293	4.302	17.861	12.464	−4.083
	GDP2	5.159	10.399	19.066	21.633	26.034	32.233	43.85	29.615
	GDP3	3.449	4.776	9.516	8.173	9.469	13.221	12.258	21.453
W		1.0031	1.0028	1.0081	1.0086	1.004	1.0105	1.0011	1.0069
D_{ij}	GDP1	−1.15	1.224	1.969	−3.07	3.6	2.341	−3.404	2.807
	GDP2	2.678	14.409	3.758	−2.488	3.218	6.75	20.826	−30.634
	GDP3	3.035	0.494	−2.644	−0.083	−1.286	−1.924	5.429	9.524
U		1.0622	1.0086	0.9835	1.0127	1.0123	1.0321	0.9716	1.027
L		1.0655	1.0114	0.9915	1.0214	1.0164	1.0429	0.9727	1.0341

Table 4 indicates that during 2002 to 2003, 2005 to 2007, 2009, the regional economic growth is faster than the overall level of Sichuan province; the growth rate is lower than the overall level of Sichuan province in 2004 and 2008. The result from the Chart tells that the main reason is a decline in the industrial sector competitiveness ($U04 = 0.9835 < 1$, $U08 = 0.9716 < 1$). During 2002 to 2009, the effective index of industrial result W is always greater than 1 which indicates the industries with potential and high-speed increase are the majority in Leshan economy. Moreover, regional overall economic structure is fine and the structure makes a great contribution to economic growth while secondary and tertiary industries are in the proportion of the optimization process step by step. During 2002 to 2003, 2005 to 2007, 2009, the effect of regional competition index U is greater than 1 which indicates the overall increasing trend of every industrial department in Leshan is great. What's more, the competitiveness of secondary industry is more obvious while the primary industry and tertiary industry cannot provide support for the regional overall competitiveness, so that the overall competitive advantages are not clear enough. When the relative competitiveness of secondary industry is on decrease, the tertiary industry is unable to provide effective support. In this situation, the secondary industry makes a decisive effect on the overall competitiveness which means this region depends on the secondary industry too much.

Table 5. 2010–2015 The SSM prediction of three industries of Leshan city

Index	Leshan	2010	2011	2012	2013	2014	2015
P_{ij}	GDP1	20.446	13.502	15.463	17.703	20.26	23.178
	GDP2	55.822	61.718	73.567	87.624	104.291	124.038
	GDP3	18.011	23.898	28.195	33.268	39.259	46.332
W		1.006	1.006	1.006	1.006	1.0059	
D_{ij}	GDP1	0.222	1.159	1.316	1.493	1.695	1.924
	GDP2	32.094	5.082	6.242	7.665	9.414	11.561
	GDP3	−5.145	1.512	1.765	2.06	2.405	2.807
U		1.0071	1.0071	1.0071	1.0071	1.0071	
L		1.0131	1.0131	1.0131	1.0131	1.013	

Choose the structure deviation component P [10] as abscissa and the competitiveness deviation component D as ordinate to draw the SSM Picture of Leshan's three industries during 2010 to 2015.

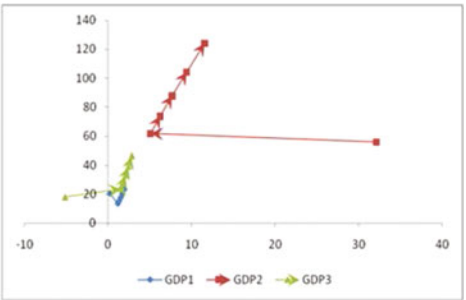


Fig. 1. 2010–2015 The SSM analysis Picture of three industries of Leshan city

Table 5 and Fig. 1 indicates: during 2010–2015, the relative growth rate (L) is greater than 1, meaning that the regional economic growth will be faster than the overall level of Sichuan province during the period of the Twelfth Five-year; W is always greater than 1 which indicates the regional overall economic structure will continue to the good development trend; further study of the structure deviation component of every industrial department find that the structure deviation component is the majority of economic part of every department. It proves that the structure of department make a great contribution to the economy growth and Leshan city already has some scale. At the same time, it is clear that compared with the secondary industry, the scale of tertiary industry has a growing difference. The effectiveness index of competitiveness shows every sector has a strong growth trend in total during the period of the Twelfth Five-year with a strong competitiveness. With the further study of the competitiveness deviation component, in addition to the obvious advantage of secondary industry, the primary industry and tertiary industry are in the average level of entire province without superiority. Still, this region depends on the secondary industry too much.

On the analysis of three industries of Leshan, we stress the further study of tertiary industry and get the SSM appraisal result of the tertiary industry during 2002–2009: the relative growth rate fluctuated, that is to say, there was no obvious advantage in its economic growth compared with the level of Sichuan province. The effect coefficient and the effect of regional competition index of every department are fluctuated too which proves that the advantages in the economic structure and competitiveness of tertiary industry are not obvious enough. With further study, we find that the competitiveness deviation component of tourism are negative generally which indicates the relative competitiveness of Leshan's tourism has not formed yet; there is a trend that the structure component of transportation industry is optimized step by step although the scale need to be further improved. Only in that way, an obvious economic effect will exists. From Table 6 and Fig. 2, the speed of growth of Leshan's tertiary industry will be slower than the level of Sichuan province in the next 5 years. We find out the result from the analysis of effect coefficient and the effect of regional competition index. The reasons are two: first, the departments in tertiary industry with high growth are minority; then, there is a non-performing growth trend of

Table 6. 2010–2015 The SSM prediction of tertiary industry of Leshan city

Index	Leshan	2010	2011	2012	2013	2014	2015
P_{ij}	GDP31	28.113	28.723	33.919	40.059	47.315	55.892
	GDP32	2.92	1.572	1.859	2.197	2.596	3.066
	SXT	26.762	34.073	40.209	47.475	56.085	66.289
W		0.995	0.9941	0.9931	0.9922	0.9913	
D_{ij}	GDP31	3.747	−2.183	−2.581	−3.051	−3.607	−4.264
	GDP32	−1.302	2.782	3.278	3.863	4.552	5.364
	SXT	−2.206	−0.648	−0.761	−0.895	−1.051	−1.236
U		0.99993	0.99992	0.99991	0.9999	0.99989	
L		0.9949	0.994	0.9931	0.9921	0.9912	

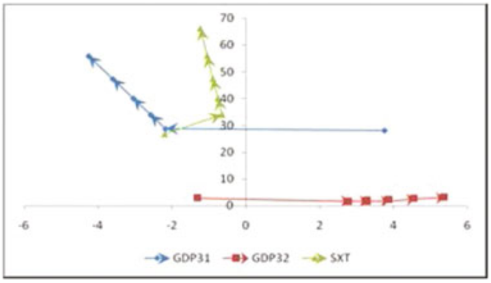


Fig. 2. 2010–2015 The SSM analysis Picture of three industries of Leshan city

every industrial department. With further analysis, the competitiveness deviation component of tourism and business are negative mainly which don't reach the average level of Sichuan province with a tendency of decrease year by year. By the study of structure deviation component, transportation and storage postal service make less contribution to economic growth with low level of structure and small scale.

4 Conclusion and Suggestion

Based on shift-share model to evaluate the industry development, we can predict the industrial development trend effectively with the help of GM(1.1) model. In this way, it can reflect the possible problems in the industrial development in the future more objectively and clearly which is convenient for related departments to plan ahead and formulate the specific targeted policies or measures. On the above analysis result, we suggest related departments to make some plans as follows:

(1) Make an accurate development orientation. It should further strengthen the development orientation of Leshan to be an important growth pole of economic development in Sichuan, a secondary transportation hub in the western, a southern core of Chengdu economic zone and an international tourist destination. The Twelfth Five-year Plan should be viewed as a period with breakthrough. Guided by the theory of low-carbon economy development, we can create a pattern of "low energy consumption, low emission, high benefit". If we speed up the pace of the adjustment of industrial structure, the industrial structure of low carbon, cycle, ecological and efficiency will formed quickly.

(2) Consolidate the leading status of secondary industry. Following with the theory that "following a new industrialization approach, developing modern industry system, promoting the combination of informatization and industrialization", we should set a core strategy of economic development in Leshan

which will promote the industrialization. Develop high-tech industry such as polysilicon, electronic information and biomedicine who have high scientific and technological content, low energy resources consumption and less pollution; Promote traditional superior industry and increase technology content; Develop low carbon economy and promote clean production which stick to the path to a new industrial city with high content of science and technology, good economic benefit, low resource consumption, less environment pollution and human capital advantage.

(3) Strengthen the competitiveness of tertiary industry. For the great increase in the proportion of the tertiary industry in the regional GDP, promote and optimize all kinds of service industries; Give full play to the unique and rich ecological tourism resources by developing ecological tourism; blossom the producer services such as information industry, green Trade and financial industry to build a new service system. Accelerate the implementation of "facilitate thousand billion transportation project and the construction of two flights four railways eight highways and one transportation hub" to build a fully functional integrated transportation system. Only in this way, the transport strength of Leshan will be promoted and the modern logistics services will be developed.

(4) Make full use of the resources rationally and promote tourism development of Leshan. Protect and make good use of the world cultural, natural heritage and bring this city's superiority into full play with famous mountain, famous Buddha, famous city and famous person. Prominent the three keys, like Buddhist culture, natural ecological tourism development and the construction management of scenic area. Improve the quality of ecological environment and the cultural taste of the city which improve the function of urban tourism. It's obvious that above actions will help speed up the process of the construction as a tourism city for Leshan.

Acknowledgements. This work was supported by the Humanities and Social Sciences Foundation ot the Ministry of Education (Grant No. 16YJC30089), and the Sichuan Center for science and Enterprise Development Research (Grant No. Xq16C04)

References

1. Artige L, Neuss L (2014) A new shift-share method. Growth Change 45(4):667–683
2. Fang L (2008) The shift-share analysis of jilin deviate from the industry's competitiveness. Master's thesis, Northeast Normal University
3. Statistics Bureau of Leshan City (2015) Leshan statistical yearbook
4. Lin CT, Yang SY (2004) Forecast of the output value of taiwan's ic industry using the grey forecasting model. Int J Comput Appl Technol 19(1):23–27
5. Liu YP (2006) Regional economics shift-share analysis of Jiangxi Province. Sci Mosaic 3:19–20

6. Otsuka A (2017) Determinants of regional energy demand: dynamic shift-share analysis. In: Regional Energy Demand and Energy Efficiency in Japan. Springer, pp 23–40
7. Shi J (2004) An analysis of the impact of sars on china. Digest Manage Sci 115:380–383
8. Statistics Bureau of Leshan City (2016) http://www.sc.stats.gov.cn
9. Wu JA, Zhao X (2009) Application of the spatial shift-share analysis. Stat Res 21:64–68
10. Zhang X, Jiao B, Jun LI (2012) Effect of component deviation on structure and infrared radiation properties of cordierite. Kuei Suan Jen Hsueh Pao/J Chin Ceram Soc 40(3):357–361
11. Zhao JH, Chen JQ, Guo LZ (2007) Application of excel in prediction and model checking of dynamic gm(1,1). J Qingdao Technol Univ 18:26–30
12. Zhou R, Li J (2015) The optimizing whitening differential equation of gm(1,1) based non-homogenous exponential series. J Math Comput Sci 5(1):50–55

Scheduling Problem for Allocating Worker with Class-Type Skill in JSP by Hybrid Genetic Algorithm

Kenichi Ida[1(✉)], Daiki Takano[1], and Mitsuo Gen[2]

[1] Graduate School of Engineering,
Maebashi Institute of Technology, Maebashi, Japan
ida@maebashi-it.ac.jp

[2] Fuzzy Logic Systems Institute and Tokyo University of Science, Tokyo, Japan

Abstract. Scheduling in manufacturing systems is one of the most important and complex combinatorial optimization problems, where it can have a major impact on the productivity of a production process. Moreover, most of manufacturing scheduling models fall into the class of NP-hard combinatorial problems. In a real world manufacturing system, a plurality of worker who operates the machine exists, depending on the skill level by the workers for each machine and working time is different even if same work on the same machine in job-shop scheduling problem (JSP). Therefore, it is taking to account for differences in working time by the worker is scheduling problem with worker allocation. In this paper, in order to approach the more realistic model by dividing into several class workers and to determine the skill level for each machine for each class worker, we propose a new model that introduced the concept of class-type skill and demonstrate the effectiveness of the computational result by Hybrid Genetic Algorithm.

Keywords: Job-shop scheduling problem (JSP) · Scheduling problem for allocating worker (SPAW) · Class-type skill · Hybrid genetic algorithm

1 Introduction

Machine scheduling problem (MSP) arises in diverse areas such as flexible manufacturing system, production planning, computer design, logistics, communication, etc. A common feature of many of these problems is that no efficient solution algorithm is known yet for solving it to optimality in polynomial time. The classical job-shop scheduling problem (JSP) is one of the most well-known MSP models. Informally, the JSP model can be described as follows: there are a set of jobs and a set of machines. Each job consists of a chain of operations, each of which needs to be processed during an uninterrupted time period of a given length on a given machine. Each machine can process at most one operation at

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_11

a time. A schedule is an allocation of the operations to time intervals on the machines. The problem is to find a schedule of minimum length [14].

In real world manufacturing systems there are many combinatorial optimization problems (COP) imposing on more complex issues, such as complex structure, non-linear constraints, and multiple objectives to be handled simultaneously. Manufacturing scheduling is one of the important and complex COP models, where it can have a major impact on the productivity of a production process. Moreover, the COP models make the problem intractable to the traditional optimization techniques be-cause most of scheduling problems fall into the class of NP-hard combinatorial problems. In order to develop effective and efficient solution algorithms that are in a sense "good", i.e., whose computational time is small as within 3 min, or at least reasonable for NP-hard combinatorial problems met in practice, we have to consider three important issues: quality of solution, computational time and effectiveness of the nondominated solutions for multiobjective optimization problem (MOP) are very important issues [2,5,6].

In a typical job-shop scheduling problem (JSP), hitherto worker's ability has not been basically considered. Scheduling problem for allocating worker (SPAW) by introducing the concept of skill in JSP as a parameter, thereby taking into account the ability of the workers. And, it is defined as an objective function to minimize the sum of the delay time for the delivery time of the completion time of each job.

Generally JSP is known as NP-hard problem which cannot be calculated in actual SPAW [3,4]. When all the operator's skill level was 1.00, it can be considered equivalent to the JSP. In other words, SPAW can be said that the generalized problem of the JSP. So, the SPAW is also NP-hard. Recently Gen [8,9] surveyed multiobjective hybrid genetic algorithms for manufacturing scheduling. Therefore, genetic algorithm has been noted that a metaheuristic approach is as a solution to the problem of the SPAW in JSP. Then the solution methods have been proposed to solve the SPAW models [11,13,15].

In this paper, we propose a new model introduced the concept of skill value by Kitada [12] which used for the nurse scheduling problem (NSP) considering SPAW based on Iima [11]. And they used the GA algorithm as a solution of the model.

The paper is organized as follows: Sect. 2 presents Scheduling Problem for Allocating Worker (SPAW); Sect. 3 describes the detailed of SPWA with the class-type skill by hybrid genetic algorithm(HGA); Sect. 4 gives a numerical experiments results; finally, the conclusion are drawn in Sect. 5.

2 Scheduling Problem for Allocating Worker (SPAW)

2.1 Workers Skill Level

In the actual production site, there is a worker to operate the machine, and the difference in processing time of work is produced in accordance with skill level on the mechanical of each worker. We introduce a skill level as a parameter to account for the difference in processing time of the operation by the worker.

By using this parameter, each worker was placed in each machine, and it is possible to determine the processing time in consideration of the worker's abilities. In addition, we can create a more realistic schedule.

(1) Processing time in consideration of skill level

When assigning N workers $W_n(n = 1, 2, \cdots, N)$ to the I products $A_i(i = 1, 2, \cdots, I)$ using K machines $M_k(k = 1, 2, \cdots, K)$, skill level $S_n(k)$ for each worker W_n is set as the ability to operate the machine M_k.

In addition, for each product $A_i(i = 1, 2, \cdots, I)$, there are Ji operations written by each one $O_{ij}(j = 1, 2, \cdots, J_i)$ to be processed by the machine$M_{R_{ij}}(R_{ij} \in 1, 2, \cdots, K)$ which is determined in advance. The processing time given to each operation O_{ij} (processing time when the skill level is 1.00) is defined as the standard processing time PT_{ij}. The processing time pt_{ij}^n of operation O_{ij} by the worker W_n is represented by the following equation, and the actual processing time expands and contracts depending on the skill level [11]:

$$pt_{ij}^n = \frac{PT_{ij}}{S_n(R_{ij})}. \tag{1}$$

The proposed worker skill level in SPAW by Iima and Sannomiya [11] is determined in a range from 0.00 to 1.00 and are given in advance as skill level table. In addition, the value of the skill level for the worker of the machine is 0.00 cannot operate the machine. Processing time on each machine by each worker will be determined using the skill level table.

(2) Setting of the Worker Skill Level

Concept of worker's skill level is twofold. One is uniform proficiency type of skill level that has been proposed by Iima [11]. Another is variety proficiency type of skill level that has been proposed by Osawa [15].

Uniform proficiency type of skill level is different for each machine. On the other hand, variety proficiency type of skill level has a skill level that each worker is different for each machine. An example of the skill level table as shown in Fig. 1.

Uniform proficiency type of skill level

	M1	M2	M3
W1	1.00	1.00	1.00
W2	0.71	0.71	0.71
W3	0.85	0.85	0.85

Varuety proficiency type of skill level

	M1	M2	M3
W1	0.71	1.00	0.00
W2	0.90	0.83	0.78
W3	0.00	0.95	0.65

(W:worker M:machine)

Fig. 1. Setting an example of the skill level of worker

2.2 Placement of Workers

The overall schedule is divided into P periods $SP_p(p = 1, 2, \cdots, P)$ and it sets the length of one period T time. At this time, each worker working at some periods of the P number of periods. For placement of workers, the following constraints exist [11].

(1) Each worker can operate the machine only with placement period.
(2) Each worker can operate one machine only at the same time.
(3) The number of workers are placed equal to the total number of machine in each period.
(4) If the work is straddling the two periods, the rest of the work will be taken over in succession to the worker after the replacement. At this time, the interruption of work at the time of change is not taken into consideration.

In SPAW proposed by Iima [11], the entire schedule is dividing to 10 periods and the length of one period is set to 8 hours. If the worker is placed skill level of 0.00 against the machine, the worker cannot operate the machine. Thus, the scheduling becomes infeasible. In order to avoid this, Osawa and Ida [15] proposed an infeasible solution correction algorithm.

2.3 Set of Delivery Time

In this paper, it is defined as an objective function to minimize the total of the late delivery time of each job. In fact, delivery time of each job is set by consultation with the customer. However, when performing experiments with an existing instance, the customer is not present. Therefore, it is necessary to set the delivery time convenience. In this paper, we carried out the delivery time settings for each job by using a delivery time coefficient [1,16].

Delivery setting using the delivery time coefficient is a technique that is used in delivery time with JSP. The value obtained by multiplying the delivery time coefficient in total processing time of each job is set as the delivery time.

In Asano [1] and Singer and Pinedo [16] of the experiment is set to delivery time coefficient $F = 1.5$ and 1.6. In addition, it has given a late delivery penalty in the case of delay in delivery occurred in each job.

3 SPAW with the Class-Type Skill by HGA

3.1 Introduction of the Class-Type Skill

In SPAW, skill level has been set in the range of 0.00 to 1.00. Skill level setting can be considered only if processing time than standard processing time is longer. However, in actual production site, worker who can handle the work in a short period of time than the standard processing time is present. In this paper, we propose a new model, which was introduced the concept of skill value Kitada [12] was used for the NSP in SPAW that Iima [11] proposed.

Divided into several class workers, to determine the skill level for each machine in each worker for each class, considering the case that can handle jobs in a shorter time than the standard processing time.

In Kitada [12] skills value table used in the NSP, the nurse of the leader, veteran, mid-level, second year, newcomers has been divided into five classes. For this reason, divided workers into five classes in this paper, to set the range of skill level in each class. Range of skill level of each class, the leader from 1.00 to 1.40, a veteran from 0.90 to 1.30, mid-level from 0.80 to 1.20, second year from 0.60 to 1.00, newcomer from 0.30 to 0.60. To determine the skill level for each machine of the workers in each of the ranges for each class. In addition, the number of workers of each class was determined based on the number ratio of each class in the skill value table of Kitada [12].

3.2 Creating Worker Group in Consideration of Average Skill Level

In SPAW of Iima [11], workers who work in each period have been prepared in advance as the worker group, and the worker group is divided into three. However, the group who work in two period (16 hours) consecutive are present in the worker group created by the Iima [11]. This is not desirable from the viewpoint of the Labor Standards Law. If you create a worker group in advance, reduced diversity of individuals, and the possibility that initial convergence occurs is high. In this paper, we set the total number of workers is three times of the total number of machines, and three of the worker group without creating in advance, we set to be created for each individual in each trial.

The creation of the worker group to use the concept of uniform proficiency type of skill level proposed by Iima [11]. Therefore, the worker has a skill level for the worker group in each class. Skill level for the worker group of each class was set the leader 1.20, veteran 1.10, mid-level 1.00, second year 0.80, and rookie 0.50. In addition, in order to prevent bias in the class of workers for each worker group, the average skill level of the worker group became when less than 0.9 will be re-create a worker group. Showing the procedure of a worker group created in consideration of the average skill level of each period of the total number of periods P as follows:

Step 1. Set $i = 1$.
Step 2. Select one worker at random from the operator that is not arranged, and to place the selected workers in the period i.
Step 3. Exclude worker placed in Step 2 from the arrangement of the choices
Step 4. If the worker of the same number as the total machine numbers are arranged in period i, the process proceeds to Step 5. Otherwise, return to Step 2.
Step 5. If it is $i = 3$, the process advances to Step 6. Otherwise, as $i = i + 1$, and return to Step 2.
Step 6. If the periods 1 to 3 skill level average of all of the worker group was 0.9 or more, the process proceeds to Step 7. Otherwise, initialize the unplaced all workers which are disposed in the periods 1–3, and returns to Step 1

Step 7. Assign the worker group of the period from 1 to 3 in order to each period of the period from 4 to P, and exit when assigned a worker group to period P.

3.3 Decision of Machine Skill Level 0 in Consideration of the Class

In the actual production site, there is a worker who cannot operate the specific machine (skill level 0.00). To account for this, the Iima [11] and Osawa [15] was set percentage of a machine skilled level 0 in advance, and based on it, when performing experiments determines the machine each worker's skill level at random becomes 0.However, in this paper, there is more than one class to worker by introducing the Class-type skill. Therefore, when using a similar method as Iima [11] and Osawa [15], for example, leader may become greater the number of machines that cannot operate than newcomers, and it is not realistic setting. In this paper, to determine machine skill level 0 in consideration of the class of workers.

An example of skill level table of the class-type skill, including a machine of skill level 0 in consideration of the class shows in Table 1.

Table 1. Skill level table using a class type skill level

Class	M_1	M_2	M_3
Leader	1.01	1.28	1.15
Veteran	1.13	0.93	1.29
Mid-level	0	0.95	1.08
2nd year	0.66	0	0.98
Newcomer	0	0	0.5

In this paper, the leader and the veteran can operate all of the machine, and mid-level, second year, newcomer is set as the operator cannot operate the part of the machine (skill level 0). In the experiment, the proportion of the worker, mid-level 10–30%, second year 20–40%, and newcomer 40–60%, is determined at random in advance.

3.4 SPAW with the Class-Type Skill by GA

Recently, many researchers tried to adapt different metaheuristic approaches such as Genetic Algorithm (GA), Tabu Search (TS), Simulated Annealing (SA) for solving job-shop scheduling problem (JSP). Because of its inherent intractability, metaheuristic procedures are an attractive alternative (Gen [7]). To improve the efficiency for finding better solutions in searching space, some special technical local searches have been adapted in JSP. Osawa [15] reformed the traditional left shift to short the idle time, and formulated an algorithm called Eshift. Goncalves et al. proposed another technique based on the critical path to confirm the search space, and swapped the operations in critical block, and this approach can also improve the efficiency of algorithm for finding active schedule (Gen [7]).

In GA for SPAW, it represents one solution using two types of chromosomes. One is a job permutation chromosome representing a processing order for each task of a job, the other is a worker placement chromosome representing the arrangement of the worker to operate the machine at each period. Therefore, genetic operations such as crossover and mutation is carried out for each of the chromosome.

In this paper, we developed hybrid genetic algorithm (HGA) for solving scheduling problem to allocate worker with class-type skill in JSP as shown in a pseudo code in Fig. 2.

```
procedure: Hybrid Genetic Algorithm
input: problem data, GA parameters
output: the best schedule
begin
    t ← 0;                                                   // t: generation no.
    initialize P1(t) by job permutation based encoding routine;
    initialize P2(t) by worker group based encoding routine; // P(t)=[P1(t), P1(t)]: population
    check & repair all chromosome P(t) for the problem constraint;
    evaluate P(t) by decoding routine and keep best solution;
    while (not terminating condition) do
        create C(t) from P(t) by crossover routine;          // C(t): offspring
        create C(t) from P(t) by mutation routine;
        check & repair the precedence constraint for all offspring C(t);
        improve offspring C(t) by Eshift routine;
        evaluate C(t) by decoding routine and update best solution;
        reproduce P(t+1) from P(t) and C(t) by selection routine;
        t ← t + 1;
    end;
    output the best schedule
end;
```

Fig. 2. A pseudo code of HGA for SPAW with class-type skill

3.5 Initial Population

In Osawa [15] of the technique, the initial population of job permutation chromosome of each individual are randomly generated. In GA, superior solution can be obtained by repeating the genetic operations such as crossover and mutation. Therefore, it is required that to have diversity in the chromosome in the initial population stage, and eliminate the diversity operation is cause to cause initial convergence.

However, when starting the operation of GA from the chromosome of the initial population randomly generated, there is a possibility that the improvement of solution stagnates. Further, since the randomly generated, it is difficult to obtain stable and excellent solution.

In this paper, we propose the initial population generation method, including the two types of chromosomes. Type one is a chromosome to be stored in the random and priority from the early work of the delivery time, and type two is a chromosome to be stored in the random and priority from the slow work of the delivery time. Initial population is generated by these chromosomes by encoding routine is the proportion of each half. This initial population generation method, while maintaining the diversity of the initial population, it is possible to have the characteristics and trends for each chromosome. This technique, because improved based on the delivery time, which is a method that can be applied to general delivery time with JSP.

3.6 Genetic Operation

(1) Job Permutation Chromosome

Osawa [15], for the job permutation chromosome, is using the proposed 2-point crossover of Hirano [10] as a crossover method, and it uses a critical block mutation as a mutation. In this paper, we introduce the concept of delivery time to 2-point crossover of the Hirano [10], and use a crossover technique to store the large work of delivery time delay time on the left side of the chromosome. This technique, because improved based on the delivery time, which is a method that can be applied to general delivery time with JSP. In addition, mutations uses Osawa [15] same critical block mutation.

(2) Worker Placement Chromosome

Osawa [15] is using only the crossover method to worker placement chromosome. That approach uses a 2-point crossover of Hirano [10] like the job permutation chromosome. In this paper, it improved based on the 2-point crossover of Hirano [10] to worker placement chromosome. Its crossover technique has a method considering skill level and total processing time for each machine in each period.

4 Numerical Experiment

4.1 Experimental Data

Make a comparison experiments with the method (oGA) by Osawa [15] with using the proposed method (pGA) in this paper. Because the Osawa [15] approach had incorporated worker group as data, the creation of worker group is using the technique proposed in this paper.

Scheduling of instances considering variety of worker skill level, such as dealing with in this paper, does not exist as far as the authors know. Therefore, as an instance in this experiment, to use instances for JSP (10job-10machine problem la16-la20, ft10). Total machine number is 10, so the total number of workers is set 30. Upon the experiment, the worker's skill level is used a randomly generated skill level in the range specified by the class-type skill level. Worker ratio of each class will be set based on the number ratio of each class in the skill value table of Kitada [5].

If the skill level for the operator of the machine is zero, the method of Osawa [15] shown in Subsect. 3.3 is used. In addition, delivery time of each job is set using the delivery time coefficient F, F was defined as $F = 1.3$ by preliminary experiment.

Various parameters for the GA are set, the number of attempts 50 times, the population size 100, crossover probability 0.8, and mutation probability 0.2. Termination condition is when the number of evaluation individual has reached one million individuals, or delivery delay time total was set to when it becomes 0. Machine specifications are, Intel Core i5-4430 3.00 GHz, memory 8.00 GB. Experiments carried out in Microsoft Visual C++ 2010 on PC.

Performing a comparison by the minimum total delivery delay time (Best) in all the attempts and the average value of the minimum total delivery delay time (Average) obtained in each trial for each instance. The results were as shown in Table 2 (number in parentheses represents the number of times when the best solution is obtained more than once).

Table 2. Experimental result

Instance	Best		Average	
	oGA	pGA	oGA	pGA
la16	60	52	135	90
la17	5	1	36	17
la18	30	14	85	55
la19	26	7	77	39
la20	0(17)	0(47)	6	0

4.2 Consideration

From Table 2, significantly better solution than the method of Osawa [15] is obtained in all instances. Figure 3 shows a Gantt chart of the best solution for the most improved instance.

In particular, the instances, la18 and la19 has become delivery delay time following conventional half. In addition, it can be seen that from the experimental results of the Average, the solution accuracy of each attempt in all instances has been greatly improved. However, it is also true that still delivery delay time occurs in these instances. In particular ft10, the width of the delay in delivery compared with the other instance has become a big thing. This is thought to be due to features of ft10. ft10 is biased machine to be used in the processing of the first few of the work of the job. Thus, even if most job can protect delivery time, jobs that delivery delay time projecting to occur are present. Therefore, it is required improvement inconsideration of the characteristics of the instances.

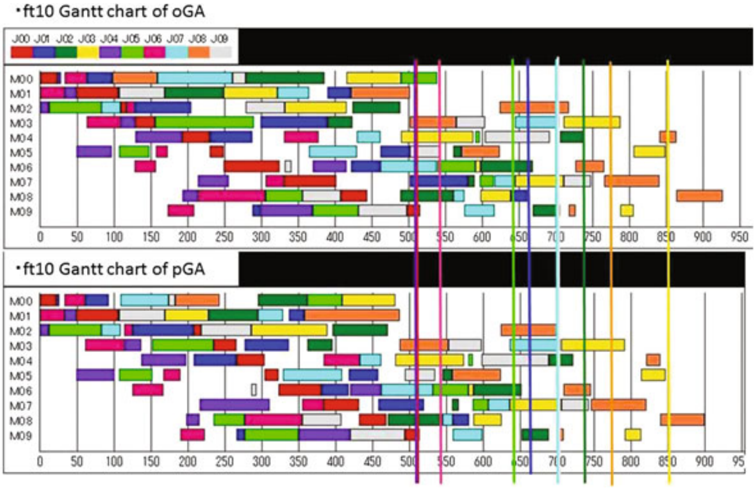


Fig. 3. ft10 Gantt chart b oGA and pGA

5 Conclusion

In this paper, scheduling problem for allocating worker (SPAW) with class-type skill in JSP expanded to the more realistic model by using Iima [11] and improvements were made of various types of algorithm using Osawa [15]. As a result, it was able to reduce the delivery delay time in all the instances that have occurred of delivery delay, and solving accuracy was able to present a schedule improved. Further, by introducing the class-type skill, divided into several class workers, to determine the skill level for each machine in each worker for each class, considering the case that can handle jobs in a shorter time than the standard processing time, and proposed SPAW of Iima [11] could be bought close to realistic models.

However, if the product had been completed earlier than the delivery time, in fact, there is a need to store the product until the delivery time. This leads to an extra cost, so it is necessary to consider in order to create a realistic schedule. The same applies to the total processing time. Therefore, after strictly to defend delivery time, it is necessary to introduce the inventory control as the second target and the total processing time as the third target.

Acknowledgements. This research work is supported by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS): No. 15K00357.

References

1. Mand AK, Hatanaka OH (2003) A heuristic algorithm for job-shop scheduling to minimize total weighted tardiness. J Jpn Ind Manage Assoc 51(3):245–253 (In Japanese)
2. Cheng R, Gen M, Tsujimura Y (1996) A tutorial survey of job-shop scheduling problems using genetic algorithmsłi: representation. Comput Ind Eng 30(4): 983–997

3. Cheng R, Gen M, Tsujimura Y (1999) A tutorial survey of job-shop scheduling problems using genetic algorithms: Part II. hybrid genetic search strategies. Comput Ind Eng 37(1C2):51–55
4. Gen M, Cheng R (1997) Genetic Algorithms and Engineering Design. Wiley, New York
5. Gen M, Lin L (2014) Multiobjective evolutionary algorithm for manufacturing scheduling problems: state-of-the-art survey. J Intell Manufact 25(5):849–866
6. Gen M, Tsujimura Y, Kubota E (1994) Solving job-shop scheduling problems by genetic algorithm. In: IEEE International Conference on Systems, Man, and Cybernetics, 1994, Humans, Information and Technology, vol 2, pp 1577–1582
7. Gen M, Cheng R, Lin L (2008) Network models and optimization: Multiobjective genetic algorithm approach
8. Gen M, Lin L, Zhang W (2015) Multiobjective Hybrid Genetic Algorithms for Manufacturing Scheduling: Part I Models and Algorithms. Advances in Intelligent Systems and Computing. Springer, Heidelberg
9. Gen M, Zhang W, Lin L (2015) Multiobjective Hybrid Genetic Algorithms for Manufacturing Scheduling: Part II Case Studies of HDD and TFT-LCD. Advances in Intelligent Systems and Computing. Springer, Heidelberg
10. Hirano H (1995) Genetic algorithm with cluster averaging method for solving job-shop scheduling sroblems. J Jpn Soc Artif Intell 27(5):769–777 (In Japanese)
11. Iima H, Sannomiya N (2002) Proposition of module type genetic algorithm and its application to modified scheduling problems with worker allocation. IEEJ Trans Electron Inf Syst 122C(3):409–416 (In Japanese)
12. Kitada M (2014) A study on a nurse scheduling system for unexpected absence of nurses. PhD thesis, Graduate School of Osaka Prefecture University, (In Japanese)
13. Kurashima Y (2013) Scheduling problem with worker allocation for FSP. Master's thesis, Maebashi Institute of Technology, (In Japanese)
14. Lawler EL, Lenstra JK (1993) Sequencing and scheduling: algorithms and complexity. Handb Ope Res Manage Sci 4(2):445–522
15. Osawa A, Ida K (2007) A solution method of scheduling problem with worker allocation by a genetic algorithm. IEEJ Trans Electron Inf Syst 127(5):755–761 (In Japanese)
16. Singer M, Pinedo M (1998) A computational study of branch and bound techniques for minimizing the total weighted tardiness in job shops. IIE Trans 30(2):109–118

The Rules Determination of Numerical Association Rule Mining Optimization by Using Combination of PSO and Cauchy Distribution

Imam Tahyudin[1,2(✉)] and Hidetaka Nambo[1]

[1] Artificial Intelligence Laboratory, Division of Electrical Engineering and Computer Science, Graduate School of Natural Science and Technology, Kanazawa University, Kanazawa, Japan
imam@blitz.ec.t.kanazawa-u.ac.jp
[2] Department of Information System, STMIK AMIKOM Purwokerto, Purwokerto, Indonesia

Abstract. One of the optimization methods to solve the numerical association rule mining problem is particle swarm optimization (PSO). This method is popularly used in various fields such as in the job scheduling problem, evaluating stock market, inferring gen regulatory networks and numerical association rule mining optimization. The weakness of the PSO is often premature for searching the optimal solution because it traps in local optima when the best particle is being searched in every iteration. Combining the PSO with Cauchy distribution for numerical association rule mining problem (PARCD) is a solution because it is robust for finding the optimal solution in a large neighborhood. The important point in this proposed method is particle representation which to know the association between one attribute to another. Therefore, this study has the aim to determinate rules of numerical association rule mining and also to calculate the multi-objective function using combination of PSO and Cauchy distribution. The results show that all of them explain every attribute to formulate the rule well. In addition, the multi-objective function value of PARCD method generally produces results which better than the previous method, MOPAR.

Keywords: PSO · Cauchy distribution · Numerical association rule mining · Particle representation · Multi-objective functions · PARCD · MOPAR

1 Introduction

The issue of numerical association rule mining is interesting to discuss. This problem already solved by many researchers with many approaches. For instance discretization [14], distribution and optimization [1,2,6,13,18]. The last approach is the interesting one because it is said as the recent solution to overcome this problem.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_12

According to [6,18] revealed that this solution can find the multi-objectives function such as support, confidence, comprehensibility, interestingness, and amplitude without transformation of numerical dataset to categorical type.

One of the methods in optimization approach is MOPAR method [6]. This method used the PSO (particle swarm optimization) for solving the numerical ARM problem. This method already developed by Tahyudin and Nambo [18] by using combination of PSO and Cauchy distribution because of the weakness of PSO which trapped in local optima when the iteration goes to infinite then the velocity particle lead to 0. This condition make the PSO does not have capability to find the optimal solution [12]. Therefore, the combination of PSO with Cauchy distribution (PARCD) is the answer for the previous method.

The important step for this method is particle representation to determine the rules. This step as the initialization process before calculating the multi-objective function. There are two approach methods, firstly is Pittsburgh method which is determine the particle as the rule set and secondly, the Michigan approach method which explains that one particle refer to one rule [6]. This research uses the second method because of being appropriate with the kind of case. Moreover, the particle representation contains three main parts, they are ACN (antecedent, consequent and none of them), the lower bound and the upper bound. Hence, the aim of this research is to determine the rules and to calculate the multi-objective function values which are support, confidence, comprehensibility and interestingness using PARCD method.

The research is organized as follows: Sect. 2 reviews the literature of recent research; Sect. 3 presents the proposed method of combining PSO and Cauchy distribution for numerical association rule mining optimization (PARCD) and the procedure to determine the particle representation; Sect. 4 gives an experiment and discussion for determining of particle representation and also the result of multi objective function; finally, the conclusion and future work are given in Sect. 5.

2 Literature Review

Minaei-Bidgoli et al. [13] demonstrated that the numerical association rule problem can be solved by discretization, distribution and optimization. The discretization is performed by partitioning and combining, clustering and fuzzy logic routines [3,5]. Then, the optimization is approached by optimized ARM [20], differential evolution [2], Genetic Algorithm (GA) [10,13,15] and Particle swarm optimization (PSO) [1,6,9]. Using PSO techniques, the numerical data can be solved to attain the important information without the discretization process [4,6] and some methods can automatically determine the minimum support and minimum confidence based on the optimal threshold [15,17,20].

On the other hand, the PSO method has the weakness that the user has to specify the number of "best rules" and the time of complexity [17] and also it is not robust when used on large data sets [7,12]. One of the ways to diminish these weakness is revealed in [12], that the combination of PSO with Cauchy has approved can rise the leverage of result because the mutation process can reach wider and appropriate to a large database. In another research work [7] the combination of approaches has the ability to optimize two-stage reentrant flexible flow shop with blocking constraints. In addition, this combination can improve the interval solution by the average of 15.60% and then the performance of this combination higher than Hybrid Genetic Algorithm (HGA) [16]. Then, this combination was used to optimize the Integration of Process Planning and Scheduling (IPPS) and the result shows the effectiveness of the proposed IPPS method and the reactive scheduling method [21]. This hybrid method was developed by [7,16] to increase the wide search space in mutation process by using Cauchy distribution. His result shows that the method can enhance the evolutionary process with the wide search space.

3 Proposed Method

3.1 The Particle Representation

The rules of numerical association rule mining by PARCD will be obtained by the particle representation procedure. This study used Michigan method which determine for every particle referring to one rule [6], for which the data set will be extracted into ACN category, based on the value of lower bound and upper bound. Antecedent is pre condition and consequent is conclusion for describing a rule. The PARCD method can classify automatically the ACN based on the optimal threshold in every rules. This concept can be showed clearly by Fig. 1.

Attribute 1			Attribute 2			……			Attribute n		
ACN_i	LB_i	UB_i	ACN_i	LB_i	UB_i				ACN_n	LB_n	UB_n

Fig. 1. The particle representation

If the optimal procedure for one rule are $0 \leq \text{ACNi} \leq 0.33$ for antecedent, $0.34 \leq \text{ACNi} \leq 0.66$ for consequent and $0.67 \leq \text{ACNi} \leq 1.00$ for none of them. For instance, see Table 1.

Table 1. Example of particle representation

Attribute	ACN_i	LBi_i	UBi_i
A	0.32	2.4	6.7
B	0.14	1.7	27.8
C	0.71	0.23	135.1
D	0.66	0.11	78.9

According to the Table 1 the attribute A and B are the antecedent and the attribute D is consequent. The attribute C is not appearing because it not includes both of them. Therefore, the rule is $AB \rightarrow D$.

3.2 Objective Design

This study uses multi-objective function which are support, confidence, comprehensibility and interestingness. The support function measures the ratio of transactions in D containing X, or Support$(x) = |x(t)|/|D|$. Witten et al. [19] explained that support is refer to coverage utility. It means that the support function can calculate the percentage of number of attributes or their combination which have the association in every data set. The support of the rule $X \rightarrow Y$ is computed using the following equation.

$$\text{Support}(X \cup Y) = \frac{|X \cup Y|}{|D|}. \tag{1}$$

The confidence measures the quality of rule based on the number of transaction of an AR in the whole dataset. In other words, the confidence describes the accuracy [19]. The rule which often emerge in every transaction is considered to have a better quality because it leads to good accuracy [6].

$$\text{Confidence}(X \cup Y) = \frac{\text{Support}(X \cup Y)}{\text{Support}(X)}. \tag{2}$$

However, this criterion does not guarantee to obtain the appropriate AR, individually. In order to gain the suitable coverage and reliability the resultant rule must also be comprehensible and interesting. According to the previous research, the smaller number of conditions in the antecedent part of a rule, the more comprehensible or understandable is that rule [8]. Hence, the comprehensibility can be measured as below

$$\text{Comprehensibility} = \frac{\log(1 + |Y|)}{\log(1 + |X \cup Y|)}, \tag{3}$$

where $|Y|$ is the number of rules in the consequent the and $|X \cup Y|$ is the total number of rules in the consequent and antecedent rules.

The interestingness criterion is used for obtaining hidden information by extracting of such surprising rules. This criterion is based on the support count of both antecedent and consequent parts [8]. The equation is shown in Eq. (4).

$$\text{Interestingness} = \left[\frac{\text{Support}(X \cup Y)}{\text{Support}(X)}\right] \times \left[\frac{\text{Support}(X \cup Y)}{\text{Support}(Y)}\right] \times \left[1 - \frac{\text{Support}(X \cup Y)}{|D|}\right]. \tag{4}$$

This formula consists of three parts. In the first part, [Support$(X \cup Y)$/Support(X)], the generation probability of the rule is computed in terms of the antecedent part of the rule. In the second part, [Support$(X \cup Y)$/Support(Y)], the generation probability of the rule is computed in terms of the consequence part of the rule. Finally, in the third part, [$1 -$ Support$(X \cup Y)/|D|$], the section Support$(X \cup Y)/|D|$ shows the probability of generating the rule according to the total number of records in the dataset ($|D|$). So, its complement, [$1-$Support $(X \cup Y)/|D|$], means the probability of not generating the rule. Therefore, a rule with a high value of support count will be considered as a less interesting rule [6].

3.3 PSO

The PSO method discovered by Kennedy and Eberhart in 1995. They are animal psychologist and electrical engineer respectively which observed the swarming behaviors in flocks of birds, schools of fish, or swarms of bees, and even human social behavior [11].

The main concept of PSO is to initialize with a group of random particles (solutions) and then to search for optima by updating generations. During all iterations, each particle is updated by following two "best" values. The first one is the best solution (fitness) has achieved so far. This value is called "pBest". The other "best" value that is tracked by the particle swarm optimizer is the best value obtained so far by any particle in the population. This best value is a global best and is called "gBest". After finding the two best values; each particle updates its corresponding velocity and position [9]. Each particle p, at some iteration t, has a position $x(t)$, and a displacement velocity $v(t)$. The personal best (pBest) and global best (gBest) positions are stored in the associated memory. The velocity and position are updated using Eqs. (5) and (6) respectively [9,12].

$$v_i^{\text{new}} = \omega v_i^{\text{old}} + c_1 \text{rand}()(p\text{Best} - x_i) + c_2 \text{rand}()(g\text{Best} - x_i), \tag{5}$$

$$x_i^{\text{new}} = x_i^{\text{old}} + v_i^{\text{new}}, \tag{6}$$

here ω is the inertia weight; $v_i$old is the particle velocity of i^{th} particle before updating; $v_i$new is the particle velocity of i^{th} particle after updating; x_i is the i^{th}, or current particle; i is the particle's number; rand() is a random number in the range (0, 1); c_1 is the individual factor; c_2 is the societal factor; pBest is the particle best; gBest is the global best Particle velocities on every dimension are clamped to a maximum velocity $V_{\max}$ [9,12].

3.4 PSO for Numerical Association Rule Mining with Cauchy Distribution (PARCD)

PARCD is the proposed method to implement for solving numerical association rule mining problem. This combination generates the best result because the weakness of PSO has solved by Cauchy distribution so it can prevent the trap of local optima in best particle to gain the global best by long jump searching using Cauchy mutation.

$$v_i(t+1) = \omega(t) v_i(t) + c_1 \text{rand}()(p\text{Best} - x_i(t)) + c_2 \text{rand}()(g\text{Best} - x_i(t)), \tag{7}$$

$$u_i(t+1) = \frac{v_i(t+1)}{\sqrt{(v_{i1}(t+1)^2 + v_{i2}(t+1)^2 + \cdots + v_{iK}(t+1))^2}}, \tag{8}$$

$$s_i(t+1) = u_i(t+1) \times \tan(\pi/2 \cdot \text{rand}[0,1)), \tag{9}$$

$$x_i(t+1) = x_i(t) + s_i(t+1). \tag{10}$$

3.5 Pseudocode and Flowchart of PARCD

This pseudocode and flowchart depict that the algorithm begins by initializing the velocity vector and position randomly then it calculates the multi-objective functions as the current fitness. After that it runs by looping iteration to looking for pBest until it finds gBest value as the optimal solution.

Pseudocode of PSO [7] is shown as following, and flowchart is as Fig. 2:

Procedure: Combination of PSO and Cauchy distribution for Numerical ARM
Input: PARCD parameters
Output: Multi-objective results
Begin
$t \leftarrow 0$
initialize $x_i(t)$ by encoding routine;
calculate the multi objectives for all particles // current fitness
evaluate $x_i(t)$ by decoding routine and keep the best solution;
while (not terminating condition) do
  for each particle xi in swarm do
    update velocity $v_i(t+1)$ // using (7)
    update position $x_i(t+1)$ //using (10)
    calculate $u_i(t+1)$ and $s_i(t+1)$ //By Cauchy dist. using (8) and (9)
    evaluate $x_i(t+1)$ //using (10)
    if $f\big(x_i(t+1)\big) < f\big(pBest_i(t)\big)$ then // $pBest_i(t)$: historical best position
      update $pBest_i(t) = x_i(t+1)$; //update the best local position
end;
  $gBest(t+1) = argmin\{f(pBest\ k(t)), gBest(t)\}$ //update the global best position
  $t \leftarrow t+1$,
  end;
  output: the best solution $gBest$;
end;

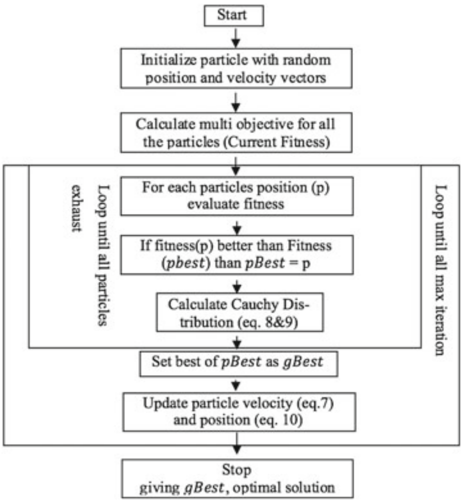


Fig. 2. Flowchart of PSO

4 Experiments and Discussion

4.1 Experimental Setup

This experiment uses benchmark datasets from Bilkent University Function Approximation Repository. There are five datasets which are used, Quake, Basketball, Body fat, pollution and Bolt (Table 2). This experiment is conducted on Intel Core i5, 8 GB main memory, running by Windows 7 and the algorithms are processed by using Matlab software.

We set up the variable of some values in the proposed algorithm such as parameter, population size, external repository size, number of iteration, the value of C_1 and C_2, W, velocity limit and xRank. They are average, 40, 100, 2000, 2, 0.63, 3.83, and 13.33 respectively. This parameters are referred according to the previous research of [6] reported in Table 3.

Table 2. Properties of the datasets

Dataset	No. of records	No. of attributes
Quake	2178	4
Basketball	96	5
Body fat	252	19
Bolt	40	8
Pollution	60	16

Table 3. The variable setup

Parameter	Population size	External repository size	Number of iteration	C_1 and C_2	W	Velocity limit	xRank
Average	40	100	2000	2	0.63	3.83	13.33

4.2 Experiments

(1) The output of particle representation
This output is generated by every running times. And then in this experiment shows the 20^{th} running time which in every running contains 2000 rules. The 20th running time is the last time so it produces the optimal solution. The results of ACN of quake data set which are the rule number, ACN and LB and UB in head title of table is shown in Table 4. According to the result in rule 1, the attributes as the antecedent are focal depth and latitude. This first rule result is same like the second rule then the last rule, rule 2000, the antecedent attributes are latitude, longitude and Richter. However, the consequent attributes are almost all none.

Table 4. ACN of quake dataset

Rule number (particle)	Antecedent, consequent, none of them (ACN)	Lower bound (LB) < attribute < Upper bound (UB)
Rule 1	Antecedent	290.028451 < Att1 < 316.467965 9.068816 < Att2 < 21.329102
	Consequence	None
Rule 2	Antecedent	290.028451 < Att1 < 316.467965 9.068816 < Att2 < 21.329102
	Consequence	None
...	...	...
...	...	...
Rule 2000	Antecedent	20.104086 < Att2 < 40.384987 33.959986 < Att3 < 71.573894 6.418029 < Att4 < 6.606453
	Consequent	None

Note: Att1 = focal depth, Att2 = latitude, Att3 = longitude, Att4 = Richter (target).

The result phenomena in Basket Ball Data set (Table 5) is almost similar to Quake data set, in which almost the rules have no consequent attributes. The first rule shows that the antecedent attributes are the number of assist per minute, height and time played. The result of second rule also shows the same as first rule, then the last rule shows that antecedent attributes are height and time played. The result of ACN in Quake and Basketball data set are interesting to study deeply, because generally the optimal solution has the antecedent and consequent attribute.

In Table 6 the result of ACN of Body Fat data set depicts the complete parameter either antecedent or consequent. In the rule 1 we see that there are eight attributes as the antecedent and three attributes as the consequents. In the rule 2 the number of antecedent and consequent attributes are the same as rule 1and in the last rule the number of antecedent and consequent attributes are six and two attributes respectively. The antecedent attributes in rule 1 are case number, Percent body fat using Siri's equation, density, age, adiposity index, chest circumference, Abdomen circumference and thigh circumference. Next, the consequent attributes are Percent body fat using Brozek's equation, height and hip circumference. Then, in rule 2 the antecedent and the consequent attributes are the same as rule 1. So that the rules 1 and 2 are if (att1, att3, att4, att5, att8, att11, att12, att14) then (att2, att7, att13). And then, in the rule 2000, the antecedent attributes are Percent body fat using Brozek's equation, Percent body fat using Siri's equation, density, height, neck circumference and knee circumference. Next, the consequent attributes are case number and weight. Therefore, the rule 2000 is if (att2, att3, att4, att7, att10, att15) then (att1, att6).

Table 7 explains the experiment result from Bolt data set which has eight attributes; run, speed, total, speed2, number2, Sens, time and T20 Bolt. Based on

Table 5. ACN of basket ball dataset

Rule number (particle)	Antecedent, consequent, none of them (ACN)	Lower bound (LB) < attribute < upper bound (UB)
Rule 1	Antecedent	0.093462 < Att1 < 0.149991 186.522680 < Att2 < 195.865376 14.763501 < Att2 < 20.155824
	Consequence	None
Rule 2	Antecedent	0.093462 < Att1 < 0.149991 186.522680 < Att2 < 195.865376 14.763501 < Att2 < 20.155824
	Consequence	None
...	...	...
...	...	...
Rule 2000	Antecedent	199.334064 < Att2 < 203.000000 24.076966 < Att3 < 39.254520
	Consequent	None

Note: Att1 = Assists_per_minute, Att2 = height, Att3 = time played, Att4 = age, Att5 = points_per_minute (target).

the table, two first rules show the same result both of antecedent and consequent. The antecedent attributes are total and time while the consequent attributes are run and speed1. Therefore, the rule is if (total, time) then (run, speed1).

The other rules do not report in this paper because they are so large, only three rules; the first, the second and the last one, rule 2000. So that, the rule 2000 shows that the antecedent attributes are run and speed 2. However, the consequent attribute is similar like quake and basket ball datasets, which is unknown. Hence. This rule can not be declared clearly because it does not have conclusion.

Table 8 depicts the result of rules from pollution dataset by particle representation PARCD method. The result from the first and the second rules are the same which the antecedent attributes are JANT, EDUC, NONW, WWDRK while the consequent attributes are PREC, JULT, OVR65, DENS and HUMID. So that, the rule is if (JANT, EDUC, NONW, WWDRK) then (PREC, JULT, OVR65, DENS and HUMID).

The rule 2000 has the ACN result which is different from the first and the second attributes. The antecedent attributes of rule 2000 are JANT, OVR65, HOUS, POOR, HC and HUMID while its consequent attributes are POPN, EDUC, DENS, NOX, SO@. The final rule is if (JANT, OVR65, HOUS, POOR, HC) then (POPN, EDUC, DENS, NOX, SO@).

(2) The output of multi-objective function
Table 9 reveals a comparison for the output of Multi-objective function among the proposed method, PARCD and MOPAR method. This table compares two objective functions which are support and confidence from both of methods in

Table 6. ACN of body fat dataset

Rule number (particle)	Antecedent, consequent, none of them (ACN)	Lower bound (LB) < attribute < upper bound (UB)
Rule 1	Antecedent	1.096724 < Att1 < 1.108900 57.988435 < Att3 < 69.574945 309.987803 < Att4 < 314.218245 55.294719 < Att5 < 66.896106 136.234441 < Att8 < 138.744999 40.927433 < Att11 < 41.562953 20.266071 < Att12 < 20.586850 22.220988 < Att14 < 23.180185
	Consequence	35.426088 < Att2 < 42.169776 113.825926 < Att7 < 122.261793 32.375620 < Att13 < 33.596051
Rule 2	Antecedent	1.096724 < Att1 < 1.108900 57.988435 < Att3 < 69.574945 309.987803 < Att4 < 314.218245 55.294719 < Att5 < 66.896106 136.234441 < Att8 < 138.744999 40.927433 < Att11 < 41.562953 20.266071 < Att12 < 20.586850 22.220988 < Att14 < 23.180185
	Consequence	35.426088 < Att2 < 42.169776 113.825926 < Att7 < 122.261793 32.375620 < Att13 < 33.596051
...	...	...
...	...	...
Rule 2000	Antecedent	12.402089 < Att2 < 18.144187 56.221481 < Att3 < 65.667791 139.024098 < Att4 < 289.982951 94.156397 < Att7 < 136.200000 57.669974 < Att10 < 87.300000 18.798957 < Att15 < 19.060978
	Consequent	1.054478 < Att1 < 1.108900 31.100000 < Att6 < 40.883823

Note: Att1 = Case Number, Att2 = Percent body fat using Brozek's equation, $457/Density - 414.2$, Att3 = Percent body fat using Siri's equation, $495/Density - 450$, Att4 = Density (gm/cm^3), Att5 = Age (years), Att6 = Weight (lbs), Att7 = Height (inches)(target), Att8 = Adiposity index = $Weight/Height^2 (kg/m^2)$, Att9 = Fat Free Weight = $(1 - fraction\,of\,body\,fat) \times Weight$, using Brozek's formula (lbs), Att10 = Neck circumference (cm), Att11 = Chest circumference (cm), Att12 = Abdomen circumference (cm) "at the umbilicus and level with the iliac crest", Att13 = Hip circumference (cm), Att14 = Thigh circumference (cm), Att15 = Knee circumference (cm), Att16 = Ankle circumference (cm), Att17 = Extended biceps circumference (cm), Att18 = Forearm circumference (cm), Att19 = Wrist circumference (cm) "distal to the styloid processes".

Table 7. ACN of bolt dataset

Rule number (particle)	Antecedent, consequent, none of them (ACN)	Lower bound (LB) < attribute < upper bound (UB)
Rule 1	Antecedent	$11.911616 < Att3 < 16.259242$ $62.782669 < Att7 < 65.562550$
	Consequence	$23.688468 < Att1 < 31.295955$ $5.928943 < Att2 < 6.000000$
Rule 2	Antecedent	$11.911616 < Att3 < 16.259242$ $62.782669 < Att7 < 65.562550$
	Consequence	$23.688468 < Att1 < 31.295955$ $5.928943 < Att2 < 6.000000$
...	...	...
...	...	...
Rule 2000	Antecedent	$13.621221 < Att1 < 29.817232$ $1.761097 < Att4 < 2.325029$
	Consequent	None

Note: Att1 = RUN, Att2 = SPEED1, Att3 = TOTAL, Att4 = SPEED2, Att5 = NUMBER2, Att6 = SENS, Att7 = TIME, Att8 = T20BOLT (target).

column side, on the row side shows the five datasets which are quake, basket ball, body fat, bolt and pollution. According to the table the global result of PARCD method is better than MOPAR method except for the support value of Quake dataset.

The highest result of support value by PARCD method is at 250.84% (Bolt dataset) which is double value than the opposite method in the same dataset. On the other hand, the lowest one is 22.97% by PARCD method which is almost less twice from the MOPAR method also in the same dataset. Next, the remaining datasets; Basket ball, Body fat and pollution, show that the PARCD has better value than MOPAR method in which the gaps are almost 30%, 60% and 10% respectively.

Then, according to the Table 9, the confidence result there is the additional value of standard deviation because it shows the stability of confidence value. Based on the the result, the confidence value generally shows that the PARCD is better than MOPAR method. It can be seen from the whole datasets from Quake to pollution data set in which the gaps are by 4%, 0.02%, 40%, 8% and 10% respectively. Interestingly, the confidence and standard deviation values from basket ball dataset have the similar values as of both of methods.

The output of comprehensibility and interestingness function can be seen in Table 10. The comprehensibility function results that the PARCD is better than MOPAR method, because there are three datasets which the values are higher. They are basket ball, body fat and pollution which the gaps are almost 125%, 130% and 40% respectively. The comprehensibility means that the results from the method are easy to understand. So, the percentage of the result can be explained that both of the methods are better than the traditional method

Table 8. ACN of pollution dataset

Rule number (particle)	Antecedent, consequent, non of them (ACN)	Lower bound (LB) < attribute < upper bound (UB)
Rule 1	Antecedent	42.431841 < Att2 < 46.441110 9.675301 < Att6 < 10.303791 24.171326 < Att9 < 27.345700 42.882070 < Att10 < 44.054696
	Consequence	21.695266 < Att1 < 22.757671 77.760994 < Att3 < 80.221960 6.698662 < Att4 < 7.071898 7436.549761 < Att8 < 7801.004046 58.816363 < Att15 < 63.240005
Rule 2	Antecedent	42.431841 < Att2 < 46.441110 9.675301 < Att6 < 10.303791 24.171326 < Att9 < 27.345700 42.882070 < Att10 < 44.054696
	Consequence	21.695266 < Att1 < 22.757671 77.760994 < Att3 < 80.221960 6.698662 < Att4 < 7.071898 7436.549761 < Att8 < 7801.004046 58.816363 < Att15 < 63.240005
...	...	...
...	...	...
Rule 2000	Antecedent	39.363260 < Att2 < 46.455909 8.721294 < Att4 < 9.206407 89.212389 < Att7 < 90.700000 21.796671 < Att11 < 23.231486 606.938956 < Att12 < 648.000000 67.768113 < Att15 < 73.000000
	Consequent	2.956662 < Att5 < 3.005372 9.450171 < Att6 < 10.068287 9345.537477 < Att8 < 9699.000000 225.061313 < Att13 < 288.274133 242.720468 < Att14 < 250.733264

Note: Att1 = PREC Average annual precipitation in inches, Att2 = JANT Average January temperature in degrees F, Att3 = JULT Average July temperature in degrees F, Att4 = OVR65 SMSA population aged 65 or older, Att5 = POPN Average household size, Att6 = EDUC Median school years completed by those over 22, Att7 = HOUS of housing units which are sound & with all facilities, Att8 = DENS Population per sq. mile in urbanized areas, 1960, Att9 = NONW non-white population in urbanized areas, 1960, Att10 = WWDRK employed in white collar occupations, Att11 = POOR poor of families with income $< \$3000$, Att12 = HC Relative hydrocarbon pollution potential, Att13 = NOX Same as nitric oxides, Att14 = SO@ Same as Sulphur dioxide, Att15 = HUMID Annual average, relative humidity at 1pm, Att16 = MORT Total age-adjusted mortality rate per 100,000.

Table 9. Output of support and confidence function

Dataset	Support (%)		Confidence (%)	
	PARCD	MOPAR	PARCD	MOPAR
Quake	22.97	46.26	86.73 ± 25.88	82.31 ± 28.91
Basket ball	61.04	32.13	92.69 ± 17.87	92.67 ± 16.65
Body fat	73.94	10.1	81.26 ± 30.67	43.59 ± 61.15
Bolt	250.84	107.29	96.88 ± 9.49	88.91 ± 9.49
Pollution	60.45	52.14	34.96 ± 43.91	23.02 ± 40.04

Table 10. Output of comprehensibility and interestingness function

Dataset	Comprehensibility (%)		Interestingness (%)	
	PARCD	MOPAR	PARCD	MOPAR
Quake	785.2 ± 37.72	786.14 ± 419.67	2.34 ± 9.30	4.67 ± 11.40
Basket ball	545.80 ± 167.74	424.65 ± 192.63	6.56 ± 21.16	4.99 ± 5.18
Body fat	333.49 ± 218.95	204.87 ± 235.46	10.61 ± 21.03	21.71 ± 9.30
Bolt	231.08 ± 168.35	271.25 ± 168.35	43.43 ± 39.68	23.70 ± 39.68
Pollution	110.63 ± 165.76	65.82 ± 130.49	9.51 ± 18.61	10.23 ± 27.88

(discretization approach). Conversely, the Interestingness result of the MOPAR is better than PARCD method which the datasets higher are quake, body fat and pollution. The gaps are about 2%, 10% and 1% respectively. Then the other data set shows that PARCD is better, using basket ball and bolt data sets by the gaps around 2% and 20% respectively.

5 Conclusions

The rules of PARCD from five data sets show the good result but there are two data sets which do not obtain the consequent, they are quake and basket ball dataset. This output is interesting to be studied continually because actually the optimal solution has the complete rule. In addition, by referring to the experiment, the output of multi objective functions which are support, confidence, comprehensibility, Interestingness, it is discovered generally that PARCD gives a better result than MOPAR method. For the future, since the problem of numerical association rule mining is still open to be improved, it will be nice to follow the research for instance by combining with other methods such as genetic algorithm or fuzzy algorithm.

Acknowledgements. This research supported by various parties. We would like to thank for scholarship program from Kanazawa University, Japan and Ministry of Research and Technology and Directorate of Higher Education (KEMENRISTEKDIKTI) and also STMIK AMIKOM Purwokerto, Indonesia. In addition, we thank for anonymous reviewers who gave input and correction for improving this research.

References

1. Alatas B, Akin E (2008) Rough particle swarm optimization and its applications in data mining. Soft Comput 12(12):1205–1218
2. Alatas B, Akin E, Karci A (2008) Modenar: multi-objective differential evolution algorithm for mining numeric association rules. Appl Soft Comput 8(1):646–656
3. Alhajj R, Kaya M (2008) Multi-objective genetic algorithms based automated clustering for fuzzy association rules mining. J Intell Inf Syst 31(3):243–264
4. Álvarez VP, Vázquez JM (2012) An evolutionary algorithm to discover quantitative association rules from huge databases without the need for an a priori discretization. Expert Syst Appl 39(1):585–593
5. Arotaritei D, Negoita MG (2003) An optimization of data mining algorithms used in fuzzy association rules. In: International Conference on Knowledge-Based and Intelligent Information and Engineering Systems. Springer, Heidelberg, pp 980–985
6. Beiranvand V, Mobasher-Kashani M, Bakar AA (2014) Multi-objective pso algorithm for mining numerical association rules without a priori discretization. Expert Syst Appl 41(9):4259–4273
7. Gen M, Lin L, Howada, (2015) Hybrid evolutionary algorithms and data mining: case studies of clustering. Proc Soc Plant Eng 2015:184–196
8. Ghosh A, Nath B (2004) Multi-objective rule mining using genetic algorithms. Inf Sci 163(1–3):123–133
9. Indira K, Kanmani S (2015) Association rule mining through adaptive parameter control in particle swarm optimization. Comput Stat 30(1):251–277
10. Kaya M (2006) Multi-objective genetic algorithm based approaches for mining optimized fuzzy association rules. Soft Comput 10(7):578–586
11. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, vol 4, pp 1942–1948
12. Li C, Liu Y et al (2007) A fast particle swarm optimization algorithm with cauchy mutation and natural selection strategy. In: International Symposium on Intelligence Computation and Applications. Springer, Heidelberg, pp 334–343
13. Minaei-Bidgoli B, Barmaki R, Nasiri M (2013) Mining numerical association rules via multi-objective genetic algorithms. Inf Sci 233:15–24
14. Narita M, Haraguchi M, Okubo Y (2002) Data abstractions for numerical attributes in data mining. Lecture Notes in Computer Science, vol 2412, pp 35–42
15. Qodmanan HR, Nasiri M, Minaei-Bidgoli B (2011) Multi objective association rule mining with genetic algorithm without specifying minimum support and minimum confidence. Expert Syst Appl 38(1):288–298
16. Sangsawang C, Sethanan K et al (2015) Metaheuristics optimization approaches for two-stage reentrant flexible flow shop with blocking constraint. Expert Syst Appl 42(5):2395–2410
17. Sarath K, Ravi V (2013) Association rule mining using binary particle swarm optimization. Eng Appl Artif Intell 26(8):1832–1840

18. Tahyudin I, Nambo H (2017) The combination of evolutionary algorithm method for numerical association rule mining optimization. In: Proceedings of the Tenth International Conference on Management Science and Engineering Management. Springer, Heidelberg, pp 13–23
19. Witten IH, Frank E (2005) Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, San Francisco
20. Yan X, Zhang C, Zhang S (2009) Genetic algorithm-based strategy for identifying association rules without specifying actual minimum support. Expert Syst Appl 36(2):3066–3076
21. Yu M, Zhang Y et al (2015) Integration of process planning and scheduling using a hybrid ga/pso algorithm. Int J Adv Manuf Technol 78(1–4):583–592

An Analytical Framework for the Conflict Coordination Mechanism of Knowledge Network

Qifeng Wei(✉)

Business School, Chengdu University of Technology,
Chengdu 610059, People's Republic of China
weiqifeng15@cdut.edu.cn

Abstract. Whether the conflict coordination mechanism works effectively directly affects the achievement of the goal of the knowledge network, and is related to both the existence of the knowledge subject itself and the knowledge network. Based on principal-agent theory and game theory, a theoretical framework of conflict coordination mechanism for knowledge network is constructed in this paper, which divides the conflict coordination mechanism into three levels: contract mechanism, self-implementation mechanism and third-party conflict coordination mechanism, then this paper makes a mathematical analysis on the theoretical framework. Results indicate that the adaptive behavior of the knowledge subjects within knowledge network evolves along with the conflict coordination mechanism framework; the conflict coordination of knowledge network contract is mainly embodied in the two aspects of benefit distribution and knowledge division; the implementation of conflict coordination relies mainly on relational contract mechanism and the trust mechanism; when the first two fail, the third party helps to relieve the relationship and make value judgments with a neutral position, in order to solve the fierce conflict.

Keywords: Knowledge network · Conflict · Coordination mechanism · Behaviors

1 Introduction

The conflicts of knowledge network refer to the state of disharmony caused by the accumulation of contradictions to a certain extent, among two or more than two knowledge subjects, due to their incompatible behaviors or goals. According to Pondy [19], the American behavioral scientist, the generation and development of conflict could be divided into five stages: latent conflict, perceived conflict, felt conflict, manifest conflict and conflict aftermath, so as to construct the embryonic of organizational conflict theory. There are many reasons for the conflict of knowledge network, that's because there are many differences between different knowledge subjects in organizational goals, management modes and organizational cultures, which will cause disagreement and lead to conflict;

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_13

The nature and scope of interdependence among knowledge subjects, are often in a dynamic reset conditions, which enables conflict inevitable; And because of the complexity of the knowledge cooperation among organizations, the dislocation of objectives will also lead to the formation of conflict; In addition, due to a changing external environment of knowledge network, inconsistencies between subjects are long-standing. Moreover, the knowledge level, knowledge structure [20], organizational routines [1], differences of values [10,11], among the knowledge subjects, also lead to various conflicts inevitably.

Whether knowledge subjects within knowledge network could learn and improve the knowledge and skills of conflict management, and accurately or effectively implement conflict management, to control the conflict without harm, is directly affecting the goal, and also it is related to subject itself and the existence of knowledge network. Therefore, in the evolution process of conflict coordination mechanism, we should objectively face various conflicts, and adopt corresponding managerial strategies to conflicts with different natures. Obviously, in the process of the generation, development and operation of the conflict mechanism of knowledge network, as the carrier of management main body and objects, knowledge subjects not only have the characteristics of cooperate with the whole team to acquire a better implementation of conflict management, but also have the initiative to actively adjust and innovate the managerial behaviors. From the perspective of long-term development practices, the behavior routines and routine variation of knowledge subjects in the conflict management, is characterized by adaptation, which is specifically the behaviors adaptation under conflicts management mechanism of knowledge network.

2 Conflict Motivation and Behavioral Framework of Knowledge Subjects

The knowledge network is composed of many knowledge subjects, and the benefit distribution scheme among the subjects has both individual rationality and collective rationality. For any subject, cooperation is better than non-cooperation, and the higher benefit can be expected. Therefore, the cooperation among subjects is a cooperative game process. Take the benefit distribution conflict as an example, Gu and Li [7] proposed the n-person cooperative game model among the members of the knowledge chain. The knowledge network consists of several knowledge chains, and the cooperation among the knowledge subjects in the knowledge network inherits the inter-organization cooperative logic of the knowledge chain, so the model is also applicable here to explain. Quan [12] argued that the formation of conflict is mainly caused by three factors: intra-member, inter-member and environment. A three-dimensional conflict model of "benefit-structure-knowledge" is an effective means to explain the conflict of knowledge network (Fig. 1).

The rational organization has the expectation of greatest interests for itself, and conflicts are generated when the benefits of each other can't be coordinated. Based on the structuration theory, the "structure" is considered to be those

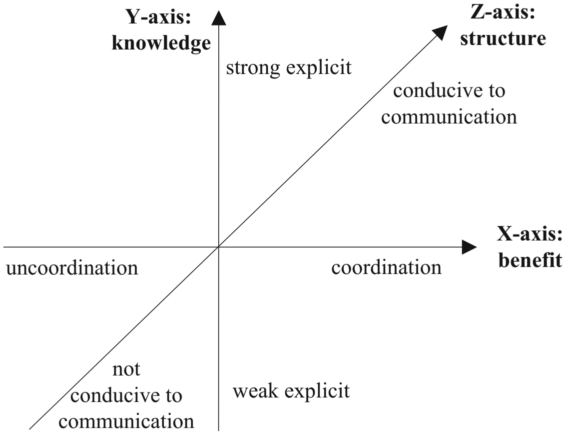


Fig. 1. "Benefit- structure - knowledge" Conflict Motivation Model

resources and institutions owned by the knowledge main body, which involves the organizational structure, organizational culture, social capital, human capital, capital, core competitive advantage and so on. The overlapping of the subjects is an important source of conflicts. Knowledge, as the basis of the operation and survival of knowledge network, determines the essence of value pursued by the knowledge subject. The development of the knowledge subject through knowledge network is to promote the knowledge creation, so as to form the knowledge advantage. Obviously, knowledge plays the central role in the conflict issue. Therefore, the conflict coordination of knowledge network also focuses on these three dimensions to operate management intervention. Wu et al. [22] argued that the problem of cooperation and conflict between knowledge subjects is actually a problem of creating and distributing income, including opportunism behavior control and benefit coordination, value creation mechanisms. The former includes a contract mechanism, a self-implementation mechanism and a third-party conflict coordination mechanism, while the latter includes a relationship adjustment mechanism and a knowledge collaboration mechanism [8]. And the framework of knowledge subject adaptive behaviors is also embedded in these mechanisms. The contract mechanism and the self-implement mechanism can effectively distribute the benefits of the knowledge network to avoid the opportunistic behavior of the knowledge subject, which is based on the coordination and interaction among the knowledge subjects. When the interaction can't solve the conflicts, then the third-party coordination mechanism is introduced to further resolve. In the mechanism of value creation, the regulation of relationship helps to improve the efficiency of cooperation among the knowledge subjects and increase the value of innovation. The mechanism of knowledge collaboration enhances the collaborative benefits of the knowledge subjects and further promotes the creation of knowledge to form the knowledge advantage.

3 The Contract Mechanism of Conflict Coordination Within Knowledge Network

According to the modern contract theory, the cooperative contract is designed to avoid opportunistic behavior and reduce transaction costs. In addition to this function, because of the characteristics of knowledge network cooperation model, the function of knowledge network contract is to promote the knowledge sharing among so as to realize the optimization of the cooperation value, which is significantly different from other types of cross-organizational consortia. For the contract design of knowledge network, the distribution of benefit and the division of knowledge are its core contents, and the conflict coordination is the important objective of those contents, that is, the rules of benefit distribution and knowledge division should effectively coordinate the possible conflicts in the process of network operation, if not, the existing benefit distribution and knowledge division rule is unscientific and unreasonable. According to the evolutionary theory, the knowledge subject has bounded rationality. Under the constraint of objective conditions, the knowledge subject tends to choose satisfaction strategy rather than the optimal strategy, which is also shown in the contract. During the operative process of knowledge network, the information asymmetry exists in the cooperation among the members, which contributes to the possibility that the information superiority side obtains the extra benefits through opportunism behaviors, thus aggravate the conflict of the knowledge network. Therefore, the design of knowledge network contract is also aimed to avoid this "rip off" problem. For the traditional contract overemphasizes partner opportunism prevention and behavior motivation, the contract of knowledge network emphasizes both cognitive and organizational coordination functions, which have a direct impact on the efficiency of knowledge network cooperation and conflict coordination. The function of knowledge-based inter-organizational cooperation contract is embodied in three aspects: motivation, coordination and cognition. The conflict coordination of knowledge network contract is mainly embodied in the two aspects of benefit distribution and knowledge division.

3.1 The Contract Mechanism of Conflict Coordination of Benefit Distribution

The design of the benefit distribution coefficient is the core issue of the conflict coordination of benefits, and the design of the contract also revolves around the core. As Wu and Gu et al. [7] stated, there are different approaches to contract optimization when considering knowledge spillovers. While the conflict of knowledge spillover is not considered, the input level of the knowledge subject is independent of the input level of the partner. In this type of contract design, the higher contribution to the overall network, the higher profit the knowledge subject should gain, which makes the incentive efficiency achieving the best, and effectively coordinate the conflict, so as to ensure the stable operation of knowledge network. Under the condition of knowledge spillover conflict, the knowledge spillover value is the important reason for the knowledge subject to choose whether to cooperate and adjust the level of knowledge input, yet the knowledge spillover loss is also the reason of the conflict formation. At this point, the input level of the knowledge subject depends on the input level of the partner, showing a positive correlation.

The study [23] revealed that, the effect of knowledge spillovers on the cooperation among knowledge subjects is a "double-edged sword", which effectively promotes cooperation but also brings conflict. On the one hand, the innovation value brought by knowledge spillovers is the driving force for deep cooperation. The stronger the knowledge absorptive capacity of knowledge subjects, the more benefits they could get from knowledge spillovers of partners, and the higher the level of knowledge input. On the other hand, because of the influence of knowledge spillover, the input level of knowledge subject is affected by the level of partner input, and the drop of knowledge input will affect the benefit from knowledge spillover. In addition, knowledge spillover may lead to the leakage of core knowledge assets, which will seriously affect the stability of knowledge network cooperation, because the huge potential value of knowledge spillover will stimulate the opportunistic behavior of the cooperators, resulting in conflict. By enhancing the level of cooperation transparency and enhancing the level of mutual trust among organizations, the observability of knowledge investment in the cooperation can be improved, which will effectively restrain opportunistic behaviors and enhance the interaction.

3.2 The Contract Mechanism of Conflict Coordination of Knowledge Division

What the knowledge division in knowledge network involves, is R&D of knowledge subject belongs to the issue of self-completion or principal-agent problem. That is, knowledge subject can do scientific research work by themselves, and make the technology development work outsourced to other subjects in the knowledge network; or both by other agency to complete. The operation mechanism of contractual mechanism of knowledge division conflict coordination includes two stages: knowledge division decision-making and multi-task principal-agent [8].

In the decision-making of knowledge division, the knowledge subject can choose independent R&D or principal-agent, and different decision-making behavior corresponds to different forms of contractual payment. Under the condition of principal-agent, in addition to receiving the entrusted payment, the entrusted entity can obtain the additional income from the technical capability growth. On the one hand, the higher the degree of R&D investment, the higher the R&D capacity gains. On the other hand, the fixed price paid by the entrusted agent represents the bargaining power in the contract conclusion, which reflects the technical capacity of the entrusted subject. From this point of view, it's clear that the higher the technical base is, the higher the additional benefits of growth will be. And because the principal does not participate in R&D, so the income is just the established distribution after the innovation success, but no additional revenue from the increase of technical capacity. According to the model of knowledge division decision-making constructed by Gu et al., the higher the coefficient of technical ability growth income, the higher the input level of the entrusted subject; the higher the payment price in contract, the higher the input level of entrusted subject. And finally the income level of principal is constructed as follows:

$$R_1^F = \frac{(1+k_1)V^2}{4}\left[\frac{\omega}{c_1^2}\right]^{\omega}\left[\frac{1-\omega}{c_2^2}\right]^{1-\omega}. \tag{1}$$

In formula (1), k_1 is the principal coefficient of technological capability growth; V is the value of the product after the success of innovation; ω is the importance level for the scientific research stage; $1-\omega$ is the level of importance for the technology development stage; the cost of the scientific research and technological development of the entrusted subject is respectively $\frac{c_1^2 e_r^2}{2}$ and $\frac{c_2^2 e_d^2}{2}$. and e_r mean e_d the level of knowledge input in two phases, $0 \le e_r \le 1$,$0 \le e_d \le 1$,$c_1^2 > 0, c_2^2 > 0$.

As the input level of the entrusted subject is positively related to its income, it can be seen that the higher the coefficient of technical capability growth, the more benefits the entrusted entity will gain. Therefore, in the decision of entrustment price, the principal will consider the technical capacity growth in the process of cooperation.

If the knowledge division uses partial delegation, that is, the knowledge subject only makes the technological development commissioned by the agent. The success of R&D depends on the degree of knowledge input and the efficiency of knowledge transfer from the stage of scientific research to technological development. Based on the model, the payment level of the principal is positively related to the importance of the development activities, that is, the more important R&D activities are, the higher the payment price will be. In addition, the higher the technical efficiency growth factor and knowledge transfer efficiency, the higher the input level of the principal. The principal's income level is as follows:

$$R_1^P = \mu^{2\omega}\frac{(1+k_2)}{2}V^2\left[\frac{\omega}{\eta^2 c_1^2}\right]^{\omega}\left[\frac{(1-\omega)^2}{2c_2^2}\right]^{1-\omega}. \tag{2}$$

In formula (2), k_2 is the coefficient of technological capability growth of subject; V is the value of the product after the success of innovation; ω is the importance level for the scientific research stage; $1-\omega$ is the level of importance for the technology development stage; μ is the knowledge transfer efficiency, $0 < \mu < 1$; η is the constant associated with R&D costs, $0 < \eta < 1$.

The choice on the form of knowledge division depends on the level of income. By comparing formulas (1) and (2), it can be seen that whether the subject of knowledge conducts the principal-agent of scientific research depends on the increase of technological capacity in R&D activities [8], the efficiency of knowledge transfer and the scientific research the importance relative to technological development.

When the knowledge subject entrust scientific research and technological development activities to other knowledge subjects, the conflict coordination becomes a multi-task principal-agent problem [9]. Based on the model of knowledge division, the model [8] considers the cost independence of R&D and the risk aversion problem of agent. In this case, the maximized deterministic equivalent income of the principal is confronted with two constraints: one is the participation constraint, the expected utility from the accepting contract is greater than or equal to the maximum expected utility when the contract is not accepted; another one is incentive compatibility constraint, when the principal is unable to follow the behaviors of the entrusted subject and the state of development, the entrusted subject always chooses a vector that maximizes its expected utility function. It can be seen that the utility behavior of any entrusted subject can only be realized through the maximization of the utility of the entrusted subject. Gu argued that the conflict coordination among knowledge subjects is affected by factors such as the profitability of entrusted subjects, the relevance of scientific research and technological development tasks, and the traceability of R&D activities.

In general, when the ability that the principal transferring the R&D results of the entrusted subject to the profit is stronger, then the degree of incentive on the entrusted subject to implement innovative behaviors is greater, which is derived from the larger profits will have more possibility for payment. When the cost of R&D is complementary, the entrusted subject will give higher motivation to the R&D task; and when the costs of R&D are substitutes for each other, if the degree of substitution is high, the degree of incentive on a single R&D task will not be too high. Similarly, if a higher incentive is given to technology development, the entrusted subject may neglect the input of scientific research activities, thus affecting the basis of product innovation; When the cost of R&D are completely independent, the incentive is related to the entrusted profit transformation ability, the marginal cost change rate, the degree of risk aversion, and the marginal cost variance and stochastic variance of another task.

4 The Self-implementation Mechanism of Conflict Coordination Within Knowledge Network

The self-implementation of conflict coordination within knowledge network is a kind of internal conflict resolution mechanism as opposed to third-party implementation. As the knowledge subject with the dominant information in the knowledge network performing "hitchhikes", "rip off" and some other acts, a series of conflicts among the organizations are triggered, although the contract effectively provides for the distribution details of innovation benefits, and to a certain extent, solve the conflict of knowledge division. However, due to the incompleteness of the contract, the prior arrangement of rights and the distribution of benefits cannot cover the content of "residual control" [25], which lead to the existence of controversy for the contract. The involvement of public relations departments, arbitration institutions, courts and other third-party conflict mediator is an optional way, but compared to the main subject of the conflict, the third-party management lacks information superiority, which leads to high management costs and diseconomy of the contractual mechanism. Therefore, it is the rational expectation of the knowledge subject to seek the maintenance and conflict coordination in the knowledge network. Dyer and Singh [4] have proposed to use self-imposed conventions to restrain opportunistic behavior. Focusing on long-term cooperative relationships in knowledge network, relational contracts among knowledge subjects can effectively reconcile possible conflicts. The relationship contract helps to make up for the short-sighted and incomplete defects of the formal contract, for instance, the conclusion of the contract can only be observed after the completion of the cooperation results, which ensures the smooth progress of knowledge cooperation. In addition, the trust among organizations is also a “dissolving agent” of the conflict. The self-implementation of the trust mechanism enhances the transparency of the knowledge subjects’ cooperation, enhances the special property investment and guarantees the openness of the communication among organizations, which is conducive to the coordination and settlement for conflict in the knowledge network.

4.1 The Relational Contract Mechanism of Cooperation and Conflict Coordination

The study of relational contract begins with the relational contract theory proposed by Macneil [15]. The theory holds that social transactions are embedded in complex relationships, thus, to study market transactions, we must make clear the various factors in social relations first, such as the specific personal relationships between contractors, the duration of the contract, the number of participants, the degree of difficulty in performing the contract, and so on. Baker and Gibbons et al. [3] argued that relational contracts are closely related to the relationship of ownership of assets, which is based on the value of future relations, it is the future relationship value based on the informal relationship. From this perspective, the purpose of the relational contract is to make the conflict subjects clear the loss of the future relationship and the loss of the direct benefit, so to change their behaviors, to promote the relaxation of the relationship to maintain the sustainable cooperation. From the perspective of contract theory, relational theory is based on the formation of asset specificity [21], whose main characteristic is “self-executing mechanism” [2]. Liu [14] even believed that the relational contract is the micro-foundation of formation and operation of the enterprise network, and the formation of the relational assets is derived from the common role of specific assets and individual social capital. Gu and Wu et al. [8] introduced the concept of relational contract to define the rent, the core enterprise and knowledge specific degree of agent organization of the relationship among the knowledge chain organizations, and analyzed the influences from discount factor and knowledge specific degree on the self-implementation of relational contract. The mechanism by which a relational contract can maintain cooperation among the contracting parties is that any breach by the other party will undermine long-term cooperative relations and thereby penalize the loss of future earnings. In order to maintain the self-implementation of the contract mechanism in the knowledge network and avoid the intervention of the third-party managers to increase the management cost, the rational knowledge subject in cooperation will adjust its behaviors from the factors closely related to the self-implementation, including relationship rent, discount factor and knowledge-specific level.

Usually, the principal and the entrusted subject cooperate to create the relationship. The more the rent is, the more the entrusted subject puts in the specific knowledge assets, the stronger the self-implementation of the relational contract is. The higher the discount factor set by the principal, the higher the input of the entrusted subject, and the stronger the self-implementation of the contract. Knowledge-specific degree refers to the relationship rent created by the unit-specific knowledge asset invested by the knowledge subject. The higher the degree of knowledge-specificity of the principal, the higher the degree of knowledge assets locked in the specific cooperative relationship with the entrusted subject, the replacement cost will be higher, so the self-implementation of relational contract is stronger; and when the principal commissioned into higher special assets, the principal must have a very high degree of knowledge in order to ensure the self-implementation of the relational contract, otherwise the subject may be subject to the implementation of opportunistic behavior to damage the benefits of the principal, thus the principal will spend more energy in the observation of the cooperative behavior of the entrusted subject.

4.2 The Trust Mechanism of Cooperation and Conflict Coordination

Network trust maintains high cooperation intensity among knowledge subjects and enhances the efficiency of knowledge transfer [18]. Trust also plays a very important role in cooperation and conflict coordination, that's because trust is the basis for establishing and maintaining intimate, cooperative, and efficient relationships and the devastation of trust due to deception and breaches can have devastating consequences for relations [24]. Trust is also a self-implementation mechanism for cooperation and conflict, based on informal rules of society, using mutual trust among organizations to make people willing to cooperate to eliminate conflicts [16,17]. The higher the degree of trust, the more significant the conflict resolution will be, on the contrary, the contradictions and conflicts will gradually increase. The role of trust in cooperation and conflict coordination among knowledge subjects in knowledge networks mainly lies in the following aspects: First, trust enhances the transparency of knowledge flow. Transparency between knowledge subjects is not high, mainly because of the vigilance on opportunistic behavior between each other, which usually lead to increased difficulty of knowledge sharing. Mutual trust can help reduce the fear of opportunistic behavior, and enhance the transparency of cooperation between organizations, thereby reducing the complexity of knowledge flow and uncertainty. Secondly, trust increases the investment in the specific assets of knowledge subjects. Specific asset investment refers to the existence of specific assets only for specific users, the knowledge subject once invest the specific asset, the bargaining power in cooperation will be reduced, which may face "rip off" behaviors. Therefore, under the condition of lack of confidence, the subject will generally choose universal asset investment rather than specific asset investment, but specific asset investment can create relational rent and help to improve performance [4]. In addition, trust helps to ensure the openness of communication among organizations. In-depth communication between organizations means that knowledge subjects have more comprehensive understanding the strategic vision, business objectives, resource allocation, financial status and other information, but deep communication between organizations must be based on a high degree of mutual trust as the prerequisite.

Trust mechanism for the cooperation between knowledge subjects and coordination of conflict, its essence is the trust game among organizations [8]. Under completely unconstrained conditions, suppose that there are two knowledge subjects, A and B, are cooperating. When they trust each other, the knowledge is fully shared. At this time, the two parties get the value of R_1 and R_2; when A trusts in B, while B is for the individual interests, bad faith and takes the opportunity to implement opportunistic behavior, so that the interests of A is damaged, $\Delta R > 0$, then the two sides get the benefits $R_1 - \Delta R$ and $R_2 + \Delta R$; If two sides do not trust each other, resulting in no knowledge sharing or knowledge sharing is in low efficiency, assuming that compared to mutual trust, the loss of benefits this time are both $2\Delta R$, so two sides respectively get $R_1 - 2\Delta R$ and $R_2 - 2\Delta R$. In the unconstrained condition, the game matrix of trust behaviors between knowledge subjects is shown in Table 1.

Table 1. Gaming matrix of trust behaviors of knowledge subjects without constraints

		Knowledge subject B	
		Trust	Distrust
Knowledge Subject A	Trust	R_1, R_2	$R_1 - \Delta R, R_2 + \Delta R$
	Distrust	$R_1 + \Delta R, R_2 - \Delta R$	$R_1 - 2\Delta R, R_2 - 2\Delta R$

As can be seen from Table 1, if knowledge subject A trusts knowledge subject B at the outset, he must bear the risk of opportunistic behavior that B is not trustworthy, and he cannot improve his condition as long as he does not change his behavioral strategy. Therefore, from the beginning, A doesn't trust B; vice versa, making the two sides into the Nash equilibrium distrust - distrust "prisoner's dilemma". The two sides of the knowledge subjects even starts from the rational assumption of self-interest purposes, the result is still detrimental to others, which is just the reason to strengthen communication to build mutual trust (Table 2).

When there are constraints, suppose knowledge subject A and knowledge subject B separately put specific assets r_1 and r_2 in cooperation. At the same

Table 2. Gaming matrix of trust behaviors of knowledge subjects with constraints

		Knowledge subject B	
		Trust	Distrust
Knowledge Subject A	Trust	R_1, R_2	$R_1 - \Delta R, R_2 + \Delta R - r_1 - p$
	Distrust	$R_1 + \Delta R - r_1 - p, R_2 - \Delta R$	$R_1 - 2\Delta R, R_2 - 2\Delta R$

time, the cooperation is influenced by informal social relations, including reputation, industry regulation and code of ethics. If someone abuses the trust to implement opportunistic behaviors, he will be punished with the loss of value p, therefore, in the presence of constraints, the payoff matrix of the trust game of knowledge subject becomes:

Compared with the income matrix in unconstrained condition, the difference exists in the losses from investment of specific assets and penalty, from a rational point of view, if $r_1 + p > \Delta R$, knowledge subject B will not choose to implement opportunistic behaviour, that is, when the extra income gained from opportunistic behaviour is less than the sum total of informal social relations punishments and specific assets input, knowledge subject B chooses to trust each other with no suspense, so that the game proceeds to the Nash equilibrium of trust-trust. Therefore, whether to knowledge subject A or B, in the consideration of the evolution of trust mechanism, negotiating the level of specific asset investment making clear social relations punishment are the main adaptive behaviours.

5 The Third-Party Conflict Management Mechanism Within Knowledge Network

The implementation of the third-party management mechanism of conflict in knowledge network is derived from the failure of the contract mechanism and the self-implementation mechanism of cooperation and conflict coordination. There is a serious disagreement among the knowledge subjects on the target and interests; it's unable to negotiate friendly. Only the third party independence institutions could help to relieve the relationship and make value judgments, to solve the fierce conflict in the knowledge network. The concept of "third party" comes from public economics, which can convey information to the conflicting parties in an impartial manner and attitude, and fairly adjudicates the conflict [13], the party is generally acted by the organization designated by government and industry associations, also, arbitration institutions and courts are often used as potential conflict resolution organizations, but its service cost is much higher.

The role of third party conflict management organization in coordinating the conflict among organizations is mainly reflected in [6,8]: to protect the benefits of the weaker subjects, to avoid the unreasonable distribution of interests, to appraise the performance of the knowledge subjects, to avoid the performance evaluation conflict of interest arising from subjective differences, and to investigate relevant conflict facts to determine the specific circumstances of the statement, so as to consider the dispute resolution. Specific functions include [13]: conflict evaluation, conflict level control, conflict resolution and conflict prevention. According to Elangovan [5], the strategy of third-party conflict control has five aspects, namely, means-control strategy (MCS), ends-control strategy (ECS), low control strategy (LCS), full control strategy (FCS), and part control strategy (PCS).

MCS mainly affect the formation process of the conflict solution, but not directly form the solution result. The final conflict resolution strategy is decided by the conflicting subject, so it is a strategy focusing on the process control, whose real form of expression is “mediation”. At this time, the adaptive behavior of knowledge subject tends to assist the management of the third party, strengthen the communication, consciously maintain the order of negotiation, discuss the solution with other conflict subjects under the witness of the third party and finally realize the reconciliation. The results of ECS is just the opposite, the strategy influences the outcome of the conflict resolution rather than the process to intervene in the dispute, directly deciding the conflict resolution, rather than trying to influence the negotiation process, the subject may positively and fully demonstrate the conflict content to a third party, the actual form of such strategies include “arbitration”, “judgment” and so on. LCS advocates using “negative” way to resolve the conflict, the third party organization only requests the conflict subject to shelve the dispute or provides the impetus to urge two sides to realize conflict self-reconciliation, rather than confront conflict itself, therefore the third-party doesn’t have the substantive involvement in conflict coordination process and the result either, For the knowledge subject, the adaptation behavior at this time is far away from conflict, shelving conflict, and achieving common development. FCS advocates the comprehensive intervention in the conflict coordination process, which resolves the conflict by influencing the way of expounding the cause of the conflict, the way of presenting the facts, and the final solution to resolve the conflict. In which the third party obtains information by asking the conflict subject about the conflict, that is, according to its own conflict resolution routines to actively search for and compile information about the conflict, and announce a conflict resolution plan to achieve the purpose, which as Elangovan stated is referred to as inquisitorial intervention or autocratic intervention, so as the parties to conflict, the knowledge subject cannot interfere in any aspect of conflict coordination, what they can only is just coordination. The last strategy PCS is a hybrid management approach, the third party implements the strategy of sharing the process of conflict-coordination and control of the strategy outcome, which is presented as: the third-party join the conflict negotiation to achieve consensus; and by facilitating the interaction and communication of the conflict subjects and arguing with the conflict subjects on the conflict event, to help think and evaluate the various options to select the best plan, recommend to convince the conflict subject to accept the preferred program to resolve conflicts, such as “arbitration after conciliation”, “collective discussion” and so on. In addition to actively cooperating with the third party to investigate and participate in collective discussion, the knowledge subject may try to find a solution to the conflict and obey the decision of the third party to eliminate conflicts. The third-party conflict intervention strategy chosen by knowledge network is mainly affected by fairness perception, satisfaction perception, effective perception and strategy implementation efficiency of knowledge subject.

6 Conclusions

The issue of conflict coordination for knowledge network, which is essential to knowledge network management, is directly influencing its effectiveness and performance. Due to this, the conflict coordination mechanism of knowledge network is studied in this paper, and an analytical framework is proposed afterwards. What's presented in this paper is: the conflict coordination of knowledge network is focusing on knowledge, benefits and structure. The contents of the cooperation and conflict coordination mechanism among knowledge subjects include contract mechanism, self-implementation mechanism and third-party conflict coordination mechanism. The framework of knowledge subject's adaptive behaviors is embedded in these mechanisms. The contract mechanism and the self-implement mechanism can effectively distribute the benefits of the knowledge network, to avoid the opportunistic behaviors of the knowledge subject, which is based on the coordination and interaction among knowledge subjects. When the interaction among the subjects cannot solve the conflict, the third-party coordination mechanism is introduced to further resolve.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (71602012), Social Science Planning Project of Sichuan Province (SC16C002), Soft Science Project of Sichuan Province (2016ZR0194), and Key Research Base Project of Philosophy and Social Sciences of Sichuan Province (CJY16-03 and Xq16C06).

References

1. Argote L, Guo JM (2016) Routines and transactive memory systems: creating, coordinating, retaining, and transferring knowledge in organizations. Res Organ Behav 36:65–84
2. Baker G, Murphy KJ (2002) Relational contracts and the theory of the firm. Q J Econ 117(1):39–84
3. Baker G, Gibbons R, Murphy KJ (1999) Informal authority in organizations. J Law Econ Organ 15(1):56–73
4. Dyer JH, Singh H (1998) The relational view: cooperative strategy and sources of interorganizational competitive advantage. Acad Manage Rev 23(4):660–679
5. Elangovan AR (1995) Managerial third-party dispute intervention: a prescriptive model of strategy selection. Acad Manage Rev 20(4):800–830
6. Fisher RJ (2016) Ronald J. Fisher: Advancing the understanding and effectiveness of third party interventions in destructive intergroup conflict. Springer

7. Gu X, Li J (2005) An n-person cooperative game analysis on the allocation of payoff among the members of knowledge chain. In: Proceedings of the 8th national conference on youth management science and system science, pp 55–62
8. Gu X, Wu S (2011) Research on conflict and conflict management among knowledge chain organizations. Sichuan University Press, Chengdu
9. Holmstrom B, Milgrom P (1991) Multi-task principal-agent analysis: Incentive contract, asset ownership and job design. J Law Econ Organ 7:24–52
10. Ibert O, Müller FC (2015) Network dynamics in constellations of cultural differences: relational distance in innovation processes in legal services and biotechnology. Res Policy 44(1):181–194
11. Kazadi K, Lievens A, Mahr D (2016) Stakeholder co-creation during the innovation process: identifying capabilities for knowledge creation among multiple stakeholders. J Bus Res 69(2):590–598
12. Li Q, Xin G (2008) The three-dimensional model of the causations of conflict among organizations in knowledge chain. Sci Sci Manage S&T 29(12):92–96 (in Chinese)
13. Liu D, Gao Y, Wu S (2010) Research on the inter-firm conflicts coordination in knowledge chain based on the third party management. J Intell 29(8):65–68
14. Liu RJ (2006) Relational contract and business networks' transition. China Ind Econ 6:91–98
15. Macneil I (1974) The many futures of contracts. South Calif Law Rev 47:691–816
16. Mayer RC, Davis JH, Schoorman FD (1995) An integrative model of organizational trust. Acad Manage Rev 20(3):709–734
17. Mcallister DJ (1995) Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. Acad Manage J 38(1):24–59
18. Miller K, Mcadam R et al (2016) Knowledge transfer in university quadruple helix ecosystems: an absorptive capacity perspective. R&D Manage 46(2):383–399
19. Purdy LR (1967) Organizational conflict: concepts and models. Adm Sci Q 13(3):296–320
20. Ranucci RA, Souder D (2015) Facilitating tacit knowledge transfer: routine compatibility, trustworthiness, and integration in m & as. J Knowl Manage 19(2):257–276
21. Williamson OE (1985) The economic institutions of capitalism: firms, markets, relational contracting. Am Polit Sci Assoc 32(2):61–75
22. Wu S (2009) Research on the stability structuring of inter-firm cooperation and conflicts in the knowledge chain. Nankai Bus Rev 12(3):54–58
23. Wu S, Gu X, Peng S (2009) Research on the knowledge spillover's influence on the stability of inter-organizational cooperation in knowledge chain. In: FSKD 2009, Sixth international conference on fuzzy systems and knowledge discovery, vol 6, pp 14–16
24. Yao Q (2011) On the trust violation and repair in the field of organizational behavioral. Nankai J 5:133–140 (in Chinese)
25. Zhang Z, Jia M, Wan D (2007) On the allocation of PPP optimal control rights from the perspective of incomplete contracts and relational contracts. Foreign Econ Manage 29(8):24–29 (in Chinese)

Flight Arrival Scheduling Optimization on Two Runways Based on IGEP Algorithm

Rui Wang, Minglei Qu(✉), and Fuzheng Wang

Business School, Sichuan University, Chengdu 610064, People's Republic of China
939774881@qq.com

Abstract. With the accelerated pace of life, more and more tourists' travel mode change from the traditional land travel into air travel. It has contributed to the rapid development of the aviation industry. But also produced some problems that troubled the airline. Among them, the flight delays problem has not been effectively addressed. The cost of flight delays is still high. This paper launches the research, analyze the cause of the cost of flight delays. For the factor of terminal area flight scheduling unreasonable to improve. Combined with the two-runway actual situation of Chengdu Shuangliu International Airport, minimize the cost of flight arrival delays, construct the model of flight arrival on two runways. At the same time, the coding method, selection strategy and fitness function of GEP are improved combined with the specific problem. Finally, IGEP and simulation are utilized to solve the practical problem. Compared with the traditional $FCFS$ rules, the cost of flight arrival delays is significantly reduced, the efficiency of flight arrival and runway utilization is improved, and the interests of airlines are guaranteed. It also shows the superiority of IGEP in addressing the issue of two-runway flight arrival.

Keywords: Flight arrival · Improved gene expression programming · Cost of flight delays · Two runways

1 Introduction

With the speeding up of life rhythm, people on the quality of travel requirements are getting higher and higher. Compared with ordinary land travel, people have a tendency to more efficient and better services' air travel. China's civil aviation passenger traffic for several years to maintain the growth rate of more than 8% [3]. In order to satisfy the needs of more passengers, major domestic airlines actively purchased aircraft, the number of aircraft China's civil aviation has been increasing for many years. At the same time, several domestic airports have built new runways for use by airlines. Two-runway airport is becoming more and more.

But the development of the aviation industry has also brought some problems cannot be ignored. Air traffic flow increases, airport traffic congestion and flight

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_14

delays are very common. According to statistics, the punctuality rate of flight in China has been lower than 80% for several years, and has a tendency to decline [4]. It is contrary to the pursuit of efficient travel. Flight delays led to some people confidence shook on the air travel, and then give up air travel, the airlines' revenue will reduce. In general, flight delays reach two hours, this flight will be a loss, the longer the delay, the higher the cost of delays. So how to control flight delays and decrease the cost of flight delays has become an urgent problem to be solved.

There are numerous reasons for flight delays, such as bad weather and aircraft faults, these accidental factors are unable to change. However, in many cases, unreasonable flight scheduling can also lead to flight delays. The data showed that nearly 40% of flight delays are related to flight scheduling. Unreasonable scheduling will aggravate flight delays, increase the cost of flight delays. On the contrary, reasonable flight scheduling can alleviate the problem of flight delays, reduce the cost of flight delays, protect the interests of airlines and improve the utilization of the runway. So how to make reasonable flight scheduling has become more and more attention by the airlines. Scholars in related fields have studied and made some achievements after noticing the importance of flight delays:

Hu and Xu [11] pointed out that flight delays and air traffic congestion due to the restriction of airport capacity, they used the ground maintenance strategy to solve the problem. Xu and Yao [19] established optimization model of multi-runway in the terminal area, and the genetic algorithm (GA) was used to solve the problem. Zhang et al. [20] constructed a dynamic multi-objective optimization model for arriving and departure flights on multiple runways. Then designed GA to solve the proposed model with the dynamic characteristics of receding horizon control strategy. Kafle and Zou [12] proposed a novel analytical-econometric approach, and built an analytical model to reveal the effects of various influencing factors on flight delays. Liang and Li [14] take flight departure time as the goal, and used improved gene expression algorithm (IGEA) to study the departure problem of single runway airport.

But the problem of two-runway flight arrival based on the cost of flight delays has not been solved. The traditional mode of flight arrival is first come first served (FCFS), without considering the impact on other flights, the advantage of this approach is completely fair between the various flights, and the scheduling is simple. It is more suitable in the early stage of China's aviation industry, the number of flights is less. With the increase in the number of flights, the shortcomings are also obvious leakage. Lack of scheduling lead to high cost of flight arrival delays, low runway utilization, cannot guarantee the interests of airlines, also cannot meet the requirements of travel time for passengers. So the traditional FCFS cannot adapt to the current flight arrival scheduling. There is an urgent need for new ways to improve the problem. Gene expression programming (GEP) is a new adaptive evolutionary algorithm based on biological structure and function [2]. It is evolved from GA and genetic programming (GP). It learns GA to encode chromosome into a fixed-length linear symbol string, and inherits the characteristics of GP to convert chromosomes into undefined expression trees. The genotype and phenotype were separated. Therefore, the use of genetic operators in terms of number and variety is not constrained. Meanwhile, mutation, insertion, and reorganization expands the search space of GEP. Using simple coding to address complex problems, make GEP greatly enhance the ability in solving specific problems. After the algorithm is proposed by Candida Ferreira, it has aroused great repercussions and became the focus of scholars in some fields. But the research of GEP is primarily focused on the prediction model and data mining [1,6,10,13]. Research on aviation management is still very limited. The application of GEP in flight arrival scheduling is more scarce. In view of GEP's powerful function, the problem of two-runway flight arrival scheduling based on IGEP is proposed in this paper. By using IGEP to solve the problem of two-runway flight arrival scheduling. The goal is to reduce the cost of flight delays, the corresponding model and IGEP were designed to optimize the problem. To ensure that the interests of airlines.

This paper is organized as follows. Section 1 introduce the reasons and influence of flight delays, discusses the study of flight scheduling, introduce the birth process and advantages of GEP, and IGEP was proposed to optimize the problem of two-runway flight arrival. Section 2 introduce the development of ZUUU and the specific problems of the flight arrival. Section 3 constructs the model based on the practical problem. In Sect. 4, some improvements of GEP are proposed, and describe the processes of IGEP. Section 5 application IGEP to solve practical case, and analyze the results of the problem; Sect. 6 is the conclusion and prospect.

2 Problem Description

Chengdu Shuangliu International Airport (ZUUU) is located in Sichuan Province, southwest of China. It was founded in 1938. After that, experienced several important expansion and development, and in November 1995 changed its name to "Chengdu Shuangliu International Airport". The second runway of the ZUUU was successfully completed and put into operation in 2009. It is located in the eastern part of the airport and connected to the first runway in the west. The aircraft can land independently on two runways. ZUUU plane diagram shown in Fig. 1:

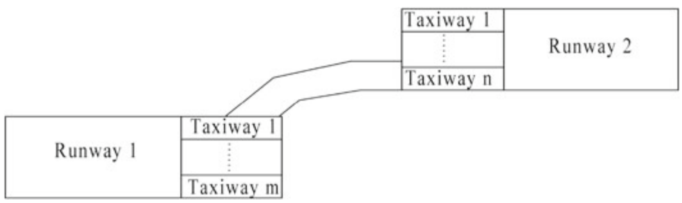


Fig. 1. ZUUU plane diagram

There are two runways in Fig. 1. Runway 1 is connected to m taxiways. Runway 2 is connected to n taxiways. For flight arrival, the runway serves as a buffer for the speed landing aircraft. Decelerated airplane slides into the taxiway, and the airplane enters the apron through the taxiway.

After the second runway put into use, the transportation capacity of ZUUU has been significantly improved, but in some cases the cost of flight arrival delays is still high. Among them, flight scheduling is the most important factor in the cost of flight arrival delays, and flight scheduling optimization mainly for the terminal area. The terminal area has two main functions. First, ensure orderly flight arrival and departure. Second, ensure safe separation and sequence between arrival flights. On the premise of flight safety, scheduling flight in the terminal area can reduce the time and cost of flight delays, improve runway utilization and ensure benefit of airlines.

Arrivaling is required after the flight reaches the terminal area. But at some point, due to the effects of bad weather and airport capacity constraints, the flight to reach the terminal area in a certain period of time will be restricted to arrival, these flights were forced to circle in the terminal area waiting to arrival. When the situation of the airport improved, these waiting flights were arranged on the runway in turn to arrival. But which flights arrival on runway 1 and which flights arrival on runway 2, as well as the sequence of flight arrival, related to the cost of flight delays. In order to minimize the total cost of flight arrival delays, these flights need to be sorted. So the problem can be transformed into the sequencing optimization problem of flight arrival. We can establish the scheduling model and design the appropriate algorithm to solve it. Flight arrival needs to consider safe wake vortex separation of the flight. Wake vortex separation is related to the aircraft type, and the cost of flight delays per unit time is not same for different aircraft types [5,7,17].

About wake vortex separation, according to the standards of international civil aviation (ICAO), all aircraft according to the maximum allowable takeoff weight is divided into three types. Heavy (H), medium (M) and light (L), and as a basis for the aircraft to set a safe wake vortex separation, ensure the safety of aircraft [8,15]. The safe wake vortex separation of ICAO is shown in Table 1.

Table 1. The ICAO safe wake vortex separation (Unit/min)

Aircraft types	Heavy (H)	Medium (M)	Light (L)
Heavy(H)	2	2	3
Medium(M)	2	2	3
Light (L)	2	2	2

In Table 1, the rows indicate the arrival flight in the front of a runway, the columns indicate the flight that is at the back on the same runway, the cross cells of row and column represent before and after the two adjacent arrival flights safe wake vortex separation on the same runway.

Table 2. The cost of flight delays per unit time for the three aircraft types

Aircraft type	The maximum allowable takeoff weight (t)	The cost of delays per unit time (yuan/min)
Heavy (H)	> 136	498
Medium (M)	7–136	236
Light (L)	≤ 7	92

About the cost of flight delays, according to the literature on the analysis of the cost of flight delays [18]. Table 2 shows the cost of flight delays per unit time for the three aircraft types, and the classification standards of aircraft type.

3 Model Construction

The total cost of flight arrival delays is minimized as the optimization goal, establish the scheduling model of flight arrival. In particular, to minimize the cost of delays in arrival flight on two runways, involved two runways, it is necessary to allocate runways to these flights, taking into account the sequence of flight arrival on the runway. The sequence of flight arrival is related to the safe wake vortex separation. The type of aircraft and the cost of flight delays per unit time is also required to consider. These factors affect the cost of flight arrival delays together.

In order to make the model more accurately describe the process of flight arrival, the following assumptions are made:

Assumption 1. The number of flights required to arrival during a certain period of time is N times, and these flights are now required to circle in the terminal area due to environmental impact and airport capacity constraints.

Assumption 2. From a certain moment, the waiting flight in the terminal area can in turn to arrival, this moment is remarked as 0.

Assumption 3. From 0 moment, the arrival flight can smoothly through the runway, taxiway into the apron. This is a continuous process. The allocation of flight on the taxiway and access to the apron does not have to be considered.

i : There are N flights circling in the terminal area waiting for arrival in a certain period of time, according to the sequence of arrival terminal area labeled as $i = 1{,}2, \cdots, N$;
j: Indicates a total of two runways $j = 1{,}2$;
k: According to the sequence of flight arrival on runway j, the flight is remarked as $k = 1{,}2, \cdots, M$, and $M \leq N$;
X_{ik}^{j}: If the $i-th$ flight in the terminal area is $k-th$ arrival on runway j, then $X_{ik}^{j} = 1$; else $X_{ik}^{j} = 0$;
u: Distinguish the aircraft types of flight. Heavy aircraft remarked as $u = 1$; medium aircraft remarked as $u = 2$; light aircraft remarked as $u = 3$;

W_u: Represents the cost of delays per unit time for the aircraft type u;
P^j_{ku}: If the $k-th$ flight arrival is the aircraft type on runway j, then $P^j_{ku} = 1$; else $P^j_{ku} = 0$;
$Q^j_{k,k-1}$: The safe wake vortex separation of adjacent flight on runway j;
T^j_k: The arrival time of the $k-th$ flight on runway j, it is also the time of flight delays;
$\triangle T^j_{k,k-1}$: The difference in arrival time between adjacent flights on runway j;
C: The total cost of flight arrival delays.

The objective function and constraints are as follows:

$$C_{\min} = \sum_{i=1}^{N}\sum_{j=1}^{2}\sum_{k=1}^{M}\sum_{u=1}^{3} P^j_{ku} Wu T^j_k X^j_{ik}, i = 1, \cdots N; j = 1, 2; k = 1, 2 \cdots, M; u = 1, 2, 3 \tag{1}$$

$$\sum_{j=1}^{2}\sum_{k=1}^{M} X^j_{ik} = 1, i = 1, 2, \cdots, N; j = 1, 2; k = 1, 2, \cdots, M \tag{2}$$

$$\sum_{i=1}^{N}\sum_{j=1}^{2}\sum_{k=1}^{M} X^j_{ik} = N, i = 1, 2, \cdots N; j = 1, 2; k = 1, 2, \cdots, M \tag{3}$$

$$\sum_{u=1}^{3} P^j_{ku} = 1, j = 1, 2; k = 1, 2, \cdots M; u = 1, 2, 3 \tag{4}$$

$$\sum_{u=1}^{3} P^j_{ku} Wu = P^j_{k1} W1 + P^j_{k2} W2 + P^j_{k3} W3, j = 1, 2; k = 1, 2, \cdots M; u = 1, 2, 3 \tag{5}$$

$$T^j_k - T^j_{k-1} = \varDelta T^j_{k,k-1}, j = 1, 2; k = 2, \cdots, M \tag{6}$$

$$\varDelta T^j_{k,k-1} \geq Q^j_{k,k-1}, j = 1, 2; k = 2, \cdots, M \tag{7}$$

$$T^j_1 = 0, j = 1, 2 \tag{8}$$

$$X^j_{ik} \in \{0,1\}, P^j_{ku} \in \{0, 1\}. \tag{9}$$

Among them, Eq. (1) is the objective function of the model, which represents the total cost of flight arrival delays; Eq. (2) shows that each flight can only arrival on one runway; Eq. (3) shows that there are N flights need to arrival at 0 moment; Eq. (4) indicates that each flight can only be one of three aircraft types; Eq. (5) shows the cost of delays per unit time of the k^{th} flight on runway j; Eq. (6) shows that the difference in arrival time between the adjacent flights on runway j is $\varDelta T^j_{k,k-1}$; Eq. (7) indicates that the difference in arrival time between adjacent flights on the same runway is not shorter than the safe wake vortex separation; Eq. (8) shows that the first flight arrival on the runway is at 0 moment and the time of delays is 0; Eq. (9) are the position and aircraft types selection variable. And then the specific algorithm is designed to solve the model.

4 Algorithm Design

FCFS has not been able to adapt to the flight arrival scheduling problem, and GEP as a new evolutionary algorithm, has a powerful function, compared with the traditional evolutionary algorithm, the ability to solve problems has been greatly enhanced. This chapter focusses on the design and process of IGEP.

4.1 IGEP

The IGEP are used to improve the problem of flight arrival on two runways and multiple taxiways. It is necessary to improve GEP according to the specific problem and model. It mainly adjusts the coding method, genetic operators design and so on of GEP, so that the algorithm is more simple and efficient in solving problems.

(1) Coding Method
Coding method around the problem, there are two runways mentioned on the model. In order to simplify the problem, it is possible to code the flight arrival on runway 1 as "1" and the flight arrival on runway 2 as "2", regardless of flight specific allocation on the taxiway. There are a total of m "1" and n "2", and N flights arrival on two runways. In accordance with the sequence of arrival, the flight arrival on runway 1 is encoded as "$A_1, A_2, \cdots, A_m$", the flight arrival on runway 2 is encoded as "$B_1, B_2, \cdots, B_n$".

(2) Mutation Operator
Mutation is the most efficient operator among all operators with the ability to modify. It can occur anywhere in the chromosome, and the chromosome structure was complete after mutation [16]. The root insertion, gene insertion, gene recombination not used in IGEP, excluding recombination and transposition. Utilising the mutation operator, the mutation probability P_m determines the point mutation in the chromosome. Since the coding used only contains the genotypes "1" and "2", the most effective mutation is to switch "1" and "2" of some genes. The gene pair represents the arrival flight on two runways, then it is easy to get the individual and population after mutation, simplify the process of GEP, increase the efficiency of solving the problem.

(3) Selection Strategy
In GEP, one of the most common selection strategies is roulette selection, in which the individual chooses according to their fitness through the roulette sampling strategy. Each individual with a round table of the disc to represent the fitness, the higher the individual fitness, the greater possibility it becomes the offspring. Instead of choosing this method, a strategy that is more suitable to solve the problem of flight scheduling is used in this paper. It preserves the individuals other than the optimal individual, and ensures that the population size remains constant.

(4) Fitness Function
The degree of fitness of the individual to the environment is expressed by the fitness, the individual fitness can be obtained by using the fitness function. The fitness function is the driving force of GEP population evolution, which directly affects the convergence speed of the algorithm and whether it can find the optimal solution [9]. In designing the fitness function, the analysis must be based on specific issues, clear objectives and accurate definition. Where the fitness function does not use the most common fitness function based on relative error. The improved fitness function is shown in Eq. (10):

$$f\left(c\right) = \frac{m+n}{C_{\min}}. \tag{10}$$

In Eq. (10), $C_{\min}$ represents the minimum total cost of flight arrival delays on two runways, m and n represents the number of arrival flights on two runways.

(5) Flight Retention Policy
The retention of the optimal flight is to retain the optimal individual and the optimal feasible individual. The optimal individual and the optimal feasible individual of the initial population are retained. Then the optimal individual and the optimal feasible individual of the first generation are compared with the individuals initially retained. If the individuals are better than the individuals originally retained, then the first generation of individuals will be retained. Otherwise individuals initially retained become the current generation. Then, in this manner successively selective substitution and retention.

4.2 IGEP Process

The IGEP steps are as follows:

Step 1. Randomly generate 50 chromosomes as the initial population;
Step 2. These chromosomes are translated into expression trees according to certain rules;
Step 3. According to the fitness function Eq. (10) to calculate the fitness of each individual;
Step 4. According to fitness value to determine whether the population to meet the requirements, if meet the requirements, the process is over; otherwise, proceed to step 5;
Step 5. Selecting the individuals according to the fitness, substitution and retention individual with selective;
Step 6. Genetic operations, the selected individuals are subjected to modified replication using genetic operators. Mutation probability is 0.05, so that the next generation to retain the new features, and then return to step 2.

The new population is subjected to the same development process described above, and the process is repeated until a satisfactory solution occurs, or iterates to a prescribed algebra, which is set to 500. According to the constructed model and the improved algorithm, the flowchart of IGEP is obtained, as shown in Fig. 2.

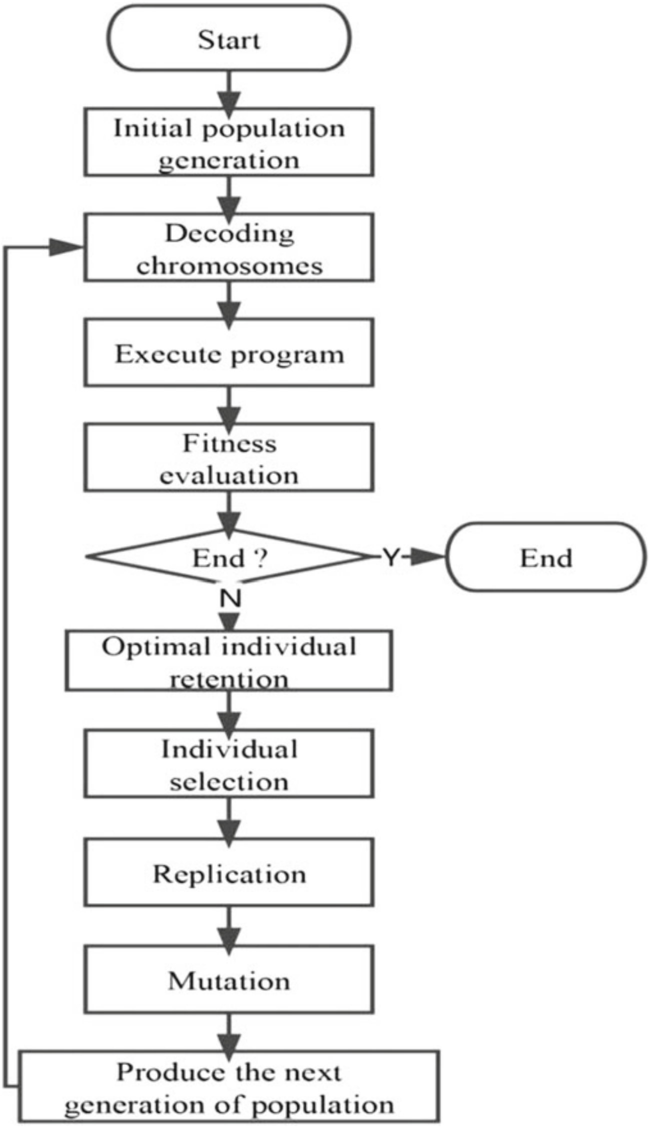


Fig. 2. Flowchart of IGEP

5 Case Simulation and Analysis

The case was developed using $C++$ and tested with multiple sets of data. The overall parameter design of IGEP is shown in Table 3.

Case taken from the ZUUU suffered a short period of time in a thunderstorm weather, resulting in flight delays, this period of time a total of 20 arrival flights were affected, leading to delay in arrival. Flights sequentially numbered 1–20

Table 3. The Parameter Design of IGEP

Parameter	Value
Prescribed algebra	500
Population size	50
Chromosome length	10
Number of genes	10
Mutation probability	0.05
Select range	100%

according to the sequence of arrival terminal area. In accordance with the FCFS rules, the odd-numbered flights arrival on runway 1 in sequence, and the even-numbered flights arrival on runway 2 in sequence, and two initial queues l_1 and l_2 are obtained according to the arrival sequence of the flight on two runways. Both queues contain 10 arrival flights, representing flights arrival sequence on runway 1 and runway 2 respectively, which is flight arrival sequence based on the FCFS rules and is also regarded as the initial flight arrival sequence L_0. According to L_0, obtained the total cost of flight arrival delays on two runways is C_0. Then the initial flight arrival sequence L_0 is optimized by IGEP according to the set parameters. When the algorithm reaches the termination condition, the optimized flight arrival sequence is L_n, and the total cost of flight arrival delays is C_n. After simulation, the flight arrival delays of the two schemes are obtained. As shown in Table 4.

Compare flight arrival delays of the two schemes in Table 4. It can be seen through optimization of IGEP. The total cost of flight arrival delays C has been significantly decreased. Compare with the traditional FCFS rules, the total cost of flight arrival delays decreased from 54,862 yuan to 30,716 yuan through the IGEP optimization. The total cost of flight arrival delays reduced 24,146 yuan, a decrease of 44.01%. The average delay cost of each flight from 2743.1 yuan down to 1535.8 yuan. The total time of flight arrival delays dropped from 43 min to 38 min, the total time decreased by 5 min, the average arrival time of each flight from 2.15 min down to 1.9 min, the flight arrival efficiency increased by 11.63%. The same flights arrival, reduced arrival time has proved that runway utilization has improved. According to the data in Table 4, the line chart of the cost of flight arrival delays on two runways are obtained, the abscissa represents the arrival sequence on the runway and the ordinate represents the total cost of flight arrival delays, as shown in Figs. 3 and 4.

As can be seen from Figs. 3 and 4, the cost of flight arrival delays on two runways has declined. The cost of flight delays on runway 1 decreased from 27406 yuan to 14590 yuan. The cost of flight delays on runway 2 decreased from 27456 yuan to 16176 yuan. Delay cost fell 12,816 yuan and 11,280 yuan respectively. The total cost of delays in airlines has significantly decreased, the income has been safeguarded, these have a positive impact on the development of airlines. At the same time, according to the trend of the line chart, we can see that when the

Table 4. The parameter design of IGEP

Flight arrival order	The initial flight arrival sequence L_0						The optimized flight arrival sequence L_n					
	Runway 1	T_1^j (min)	Delay cost (yuan)	Runway 2	T_2^j (min)	Delay cost (yuan)	Runway 1	T_1^j (min)	Delay cost (yuan)	Runway 2	T_2^j (min)	Delay cost (yuan)
1	H(1)	0	0	H(2)	0	0 H(2)	0	0	H(2)	0	0	
2	L(1)	3	276	M(2)	2	472 H(1)	2	996	H(1)	2	996	
3	M(1)	5	1456	L(2)	5	932	H(1)	4	2988	H(2)	4	2988
4	L(1)	8	2192	H(2)	7	4418	M(2)	6	4404	M(1)	6	4404
5	M(1)	10	4552	L(2)	10	5338	M(1)	8	6292	M(2)	8	6292
6	H(1)	12	10528	M(2)	12	8170	M(2)	10	8652	M(1)	10	8652
7	L(1)	15	11908	L(2)	15	9550	L(2)	13	9848	M(2)	12	11484
8	H(1)	17	20374	H(2)	17	18016	L(1)	15	11228	L(2)	15	12864
9	L(1)	20	22214	M(2)	19	22500	L(2)	17	12792	L(1)	17	14428
10	M(1)	22	27406	M(2)	21	27456	L(1)	19	14540	L(1)	19	16176
Delay cost	54862											
Delay time	43											

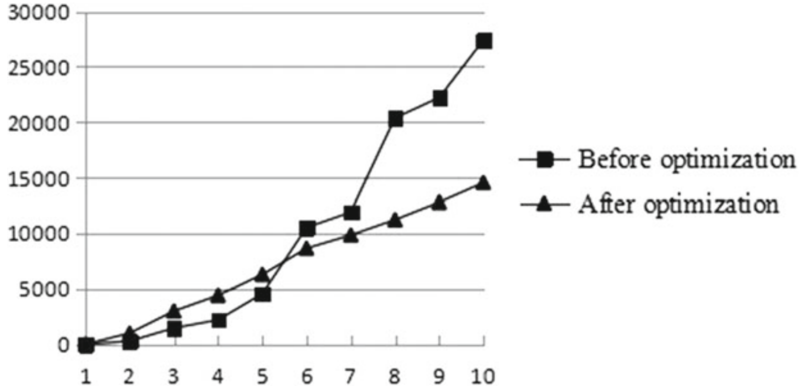


Fig. 3. The cost of flight arrival delays on runway 1

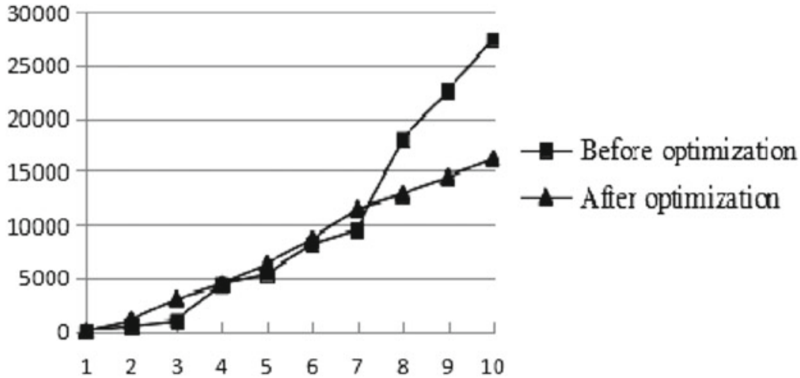


Fig. 4. The cost of flight arrival delays on runway 2

number of arrival flights reaches a certain scale, with the increasing number of arrival flights, the difference between the cost of flight arrival delays before and after optimization is greater, and the economic significance of the optimization of the flight arrival problem becomes more apparent.

6 Conclusion

The rapid development of social economy has promoted the change of passenger's travel mode. The demand for aviation industry has been growing. The development of aviation industry has also resulted in flight delays and high cost of flight delays, which impose higher demands on flight arrival scheduling. This paper studies the problem of flight arrival delays based on the results of domestic and international flight scheduling research. The total cost of flight arrival delays is minimized as the optimization goal, construct the scheduling model of flight arrival, and IGEP is designed according to the specific problem and model.

Through practical case studied of ZUUU, it was found that IGEP-optimized flight arrival scheme, compared with the traditional FCFS scheme, the cost of flight delays is significantly reduced, also reducing the arrival time of the flight, the efficiency of flight arrival and runway utilization are improved. It also showed the effectiveness of IGEP in solving the scheduling problem of flight arrival. For airlines, this approach is worth promoting. It can reduce the loss of the airlines in solving the scheduling problem of two-runway flight arrival, and ensure the interests of airlines. It also has a positive impact on the development of the aviation industry. The research work in the future can be associated with flight scheduling on the taxiway, improve the process of flight arrival. In the aircraft type, the three types can be more detailed division, accurate to each type, in order to get the conclusion more convincing.

References

1. Alavi A (2011) A robust data mining approach for formulation of geotechnical engineering systems. Eng Comput 28:242–274 (in Chinese)
2. Azamathulla H, Ahmad Z, Aminuddin A (2013) Computation of discharge through side sluice gate using gene-expression programming. Irrig Drain 62:115–119 (in Chinese)
3. Civil Aviation Administration of China (2016) The civil aviation industry in 2015 statistical bulletin [eb/ol]
4. Development Planning Department of Civil Aviation Administration of China (2016) From the statistical view of Civil Aviation. China Civil Aviation, Beijing
5. Deshpande V, Kan M (2012) Impact of airline flight schedules on flight delays. Manufact Serv Oper Manage 14:423–440 (in Chinese)
6. Divsalar M (2012) A robust data-mining approach to bankruptcy prediction. J Forecast 31:504–523 (in Chinese)
7. Dorndor U (2007) Disruption management in flight gate scheduling. Statistica Neerlandica 61:92–114 (in Chinese)
8. Drüe C (2008) Aircraft type-specific errors in amdar weather reports from commercial aircraft. J R Meteorol Soc 134:229–239 (in Chinese)
9. Fernández-Ares A (2016) Analyzing the influence of the fitness function on genetically programmed bots for a real-time strategy game. Entertainment Comput 18:15–29 (in Chinese)
10. Gandomi A (2011) A new prediction model for the load capacity of castellated steel beams. J Constr Steel Res 67:1096–1105 (in Chinese)
11. Hu M, Xu X (1994) Ground holding strategy for air traffic flow control. J Nanjing Univ Aeronaut Astronaut 26:26–30 (in Chinese)
12. Kafle N, Zou B (2016) Modeling flight delay propagation: a new analytical-econometric approach. Transp Res Part B Methodol 93:20–542 (in Chinese)
13. Karbasi M, Azamathulla H (2016) Gep to predict characteristics of a hydraulic jump over a rough bed. KSCE J Civil Eng 20:1–6 (in Chinese)
14. Liang W, Li Y (2014) Research on optimization of flight scheduling problem based on improved gene expression algorithm. Comput Technol Dev 7:5–8 (in Chinese)
15. Marques J (2016) On an analytical model of wake vortex separation of aircraft. Aeronaut J 120:1534–1565 (in Chinese)
16. Peng J (2015) A new evolutionary algorithm based on chromosome hierarchy network. Int J Comput Appl 30:183–191 (in Chinese)

17. Vasilyev I, Avella P, Boccia M (2016) A branch and cut heuristic for a runway scheduling problem. Autom Remote Control 77:1985–1993 (in Chinese)
18. Xie T (2009) Study of arrival flight scheduling optimizing based on delay cost. PhD thesis, Beijing Jiaotong University, Beijing (in Chinese)
19. Xu X, Yao Y (2004) Application of genetic algorithm to aircraft sequencing in terminal area. J Traffic Transp Eng 4:121–126 (in Chinese)
20. Zhang Q, Hu M, Zhang H (2015) Dynamic multi-objective optimization model of arrival and departure flights on multiple runways based on RHC-GA. J Traffic Transp Eng 2:012 (in Chinese)

A Study of Urban Climate Change Vulnerability Assessment Based on Catastrophe Progression Method

Yang Sun[1], Yi Lu[1](✉), and Yinghan Li[2]

[1] Business School, Sichuan University, Chengdu 610065, People's Republic of China
luyiscu@163.com

[2] Software School, Harbin University of Science and Technology, Harerbin 150022, People's Republic of China

Abstract. Climate change can make a profound impact on the entire ecosystem, and severely affects the stability of human society. The urban is the economic, political and cultural center of human society. The assessment of urban disaster vulnerability is of great significance to the disaster prevention system. Through the review of the relevant literatures on disasters vulnerability and the situations of Sichuan Province, the study establishes urban climate change actual vulnerability assessment systems. The vulnerability degrees of the five economic zones of Sichuan Province are evaluated by the improved mutation progression method based on the entropy method. The study shows that the urban disaster vulnerability of Southern Sichuan Economic Region is the most serious, and the degree of disaster vulnerability of Chengdu Economic Region is the lowest. Chengdu Economic Region should concern about the floating population. And Southern Sichuan Economic Region needs to create a good environment for enterprises to increase the number of enterprises within its jurisdiction.

Keywords: Climate change · Vulnerability · Sichuan province · Mutation progression method · Entropy method

1 Introduction

Climate change has already made a serious impact on human beings [1]. It not only makes social economy worse [3], increases the flood exposure of cities [6] and the difficulty of reservoir management [14], severely threaten the human health [11], but also destroys the diversity and the stability of ecosystem [2,9]. Therefore, climate change has been becoming an important issue related to the survival and sustainable development of human beings. To minimize the threat caused by climate change, effective measures need to be formulated.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_15

The vulnerability can comprehensively assess the impact of natural disasters [7], and provide a fundamental basis for effective adaptation to climate change [12]. And the vulnerability assessment of urban climate change disasters is an important link to enhance the ability of urban disaster response, and has important practical significance [8].

After years of development, foreign vulnerability research has acquired remarkable achievements. The Intergovernmental Panel on Climate Change (IPCC) has established a system of climate change vulnerability method, which has been applied to 24 countries [5]. The Tyndall Center constructed a national climate vulnerability assessment system [15]. China's relevant researches on the vulnerability of climate change disasters began in 1990s, which mainly were related to the impact of climate change on water resources and agricultural disaster, rarely including the regional composite system [17].

The influences caused by climate change are widespread and multi-level [11]. However, the current research focuses on the impact of climate change on the coastal cities [6], few of those studies of climate change in inland cities. In our study, Sichuan Province, an inland region of China, is regarded as the research object. The urban climate change vulnerability study of Sichuan five economical areas is based on catastrophe progression method. The result of our survey can provide objective statement of economical regions' urban vulnerability, and offer effective references to urban climate change response.

2 Method

The catastrophe progression method combines catastrophe theory and fuzzy mathematics [10]. And in order to make the sort of indicators more objective and reduce the influence of subjective factors, the evaluation indexes are ranked by the weight. The catastrophe progression method can be divided into 6 steps.

Step 1. Establish assessment systems.

The overall goal is divided into multi-level indicators. And the total number of the corresponding indicators is not more than four, which can be expressed by the parameter n.

Step 2. Identify the model of system.

$$f(x) = \begin{cases} x^3 + ax, & n = 1 \\ x^4 + ax^2 + bx, & n = 2 \\ \frac{1}{5}x^5 + \frac{1}{3}ax^3 + \frac{1}{2}bx^2 + cx, & n = 3 \\ \frac{1}{6}x^6 + \frac{1}{4}ax^4 + \frac{1}{3}bx^3 + \frac{1}{2}cx^2 + dx, & n = 4. \end{cases} \tag{1}$$

The $f(x)$ represents the potential function of the state variable x. And a, b, c represent the control variable of the state variable x.

Step 3. Data standardization.

$$x_{ij}{}^{*} = \begin{cases} x_{ij}/\max\limits_{j}(x_{ij}), & x_{ij} \in x_{ij}^{+} \\ \min\limits_{j}(x_{ij})/x_{ij}, & x_{ij} \in x_{ij}^{-}. \end{cases} \tag{2}$$

The x_{ij} means the evaluation index j of object i. The x_{ij}^{+} is a benefit attribute. The x_{ij}^{-} is a cost type attribute. The $\max_j(x_{ij})$ is maximum value of evaluation index j. The $\min_j(x_{ij})$ is minimum value of evaluation index j.

Step 4. Calculate the weight of each index by the entropy method, and rank the evaluation indexes based on the weights of evaluation indexes.

$$p_{ij} = x_{ij}^{*} \Big/ \sum_{i=1}^{m} x_{ij}^{*}, j = 1, 2, \cdots, n, \tag{3}$$

$$e_j = -\frac{1}{\ln m} \sum_{i=1}^{m} p_{ij} \ln p_{ij}, j = 1, 2, \cdots, n, \tag{4}$$

$$w_j = (1 - e_j) \Big/ \sum_{j=1}^{n} (1 - e_j), j = 1, 2, \cdots, n. \tag{5}$$

The m represents the number of evaluated objects. And n indicates the number of indicators. The value of w_j is the weight of each index.

Step 5. Confirm the normalized equation of each model.

$$x = \begin{cases} x_a = a^{1/2}, & n = 1 \\ x_a = a^{1/2}, x_b = b^{1/3}, & n = 2 \\ x_a = a^{1/2}, x_b = b^{1/3}, x_c = c^{1/4}, & n = 3 \\ x_a = a^{1/2}, x_b = b^{1/3}, x_c = c^{1/4}, x_d = d^{1/4}, & n = 4. \end{cases} \tag{6}$$

Step 6. Comprehensive evaluation.

According to the multi objective fuzzy decision theory, $T_1, T_2, \cdots, T_n$ can be regarded as a fuzzy goal for the same plan in the multi-objective situation. Then the ideal membership function is $u(x) = u_{T_1}(x) \cap u_{T_2}(x) \cap \cdots \cap u_{T_n}(x)$, of which $u_{T_i}(x)$ is the function of the fuzzy decision T_i. By unitary formula, the corresponding value of the state variable can be confirmed.

3 Case Study

Sichuan province is located in the hinterland of the China's southwest, and on the eastern of Qinghai-Tibetan plateau. Due to the unique geographical environment of the plateau, the climate change in Sichuan is very obvious. Meanwhile, Sichuan is located in the Longmen Shan Fault, in which geological disasters easily happen [13]. The sudden heavy rainfall caused by climate change led to the city water logging, and even geological disasters (e.g., landslide, debris flow). The basic situation of cities in each economic zone is shown in Fig. 1.

3.1 Construct the Index System

The vulnerability of urban disaster prevention system consists of response, sensitivity, resilience [4]. The related researches of vulnerability don't form a unified vulnerability assessment system, and the selection of vulnerability index

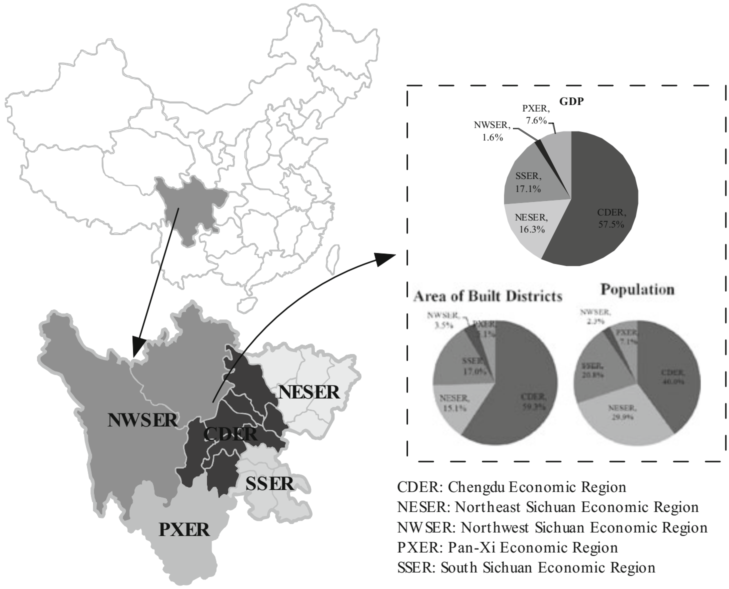


Fig. 1. Overview diagram of Sichuan Province

varies according to the real problems [16]. Sichuan Province is divided into five major economic zones, including Chengdu Economic Region (CDER), Northeast Sichuan Economic Region (NESER), Northwest Sichuan Economic Region (NWSER), Pan-Xi Economic Region (PXER), and South Sichuan Economic Region (SSER). In order to make the economic zone of the social vulnerability index objective and reasonable, the vulnerability assessment of urban climate change disaster is established based on actual situation of Sichuan (Table 1).

3.2 Assess the Urban Vulnerability of Climate Catastrophe

Our study regards Sichuan Statistical Yearbook (2015) as the source of the original data. According to Eqs. (2)–(6), the standard data and the importance of each indicator can be calculated. Table 1 shows the sorted result of each layer index.

In the assessment system, the number of each layer's indexes is not more than 4. According to the number of each level's indicators, the vulnerability of all levels' indicators can be calculated. The urban vulnerability of Sichuan economic zone is shown in Table 2. The larger the number of values in the table, the more serious the degree of vulnerability. As shown in Table 2, the disaster vulnerability of the CDER is the lowest, and the ability of disaster prevention and mitigation is the highest. The vulnerability of the SSER is the highest, whose ability of disaster prevention and reduction is the worst.

Table 1. Vulnerability assessment system and index importance of urban climate change disasters prevention system

The target layer	Criterion layer (Importance)	Index layer (Importance)	Measure (Importance)
Urban vulnerability of climate catastrophe	Capacity of coping A (0.6285)	Primary organization $A1$ (0.4811)	The number of Enterprises $A11$ (0.3190)
			The number of neighborhood committees $A12$ (0.1620)
		Traffic situation $A2$ (0.0873)	Per capita road area $A21$ (0.0579)
			Per capita car number $A22$ (0.0294)
		Awareness of disaster $A3$ (0.0404)	Per capita publicity expenditure $A31$ (0.0263)
			Per capita educational expenditure $A32$ (0.0141)
		Medical condition $A4$ (0.0197)	Per capita number of medical institutions $A41$ (0.0121)
			Per capita number of medical personnel $A42$ (0.0039)
			Per capita number of sickbeds $A43$ (0.0036)
	Sensitivity B (0.2125)	Urban status $B1$ (0.1240)	Drain pipe length $B11$ (0.1237)
			Per capita housing area $B12$ (0.0002)
			Greening rate $B13$ (0.0001)
		The structure of population $B2$ (0.0885)	Population ratio of Higher Education $B21$ (0.0499)
			Ratio of old people $B22$ (0.0223)
			Population density $B23$ (0.0062)
	Resilience C (0.1590)	Self rescue ability $C1$ (0.1304)	Institutional deposit amount $C11$ (0.1299)
			Per Capita disposable income $C12$ (0.0006)
		Social security $C2$ (0.0285)	Coverage rate of unemployment insurance $C21$ (0.0223)
			Coverage rate of medical insurance $C22$ (0.0062)

Table 2. Assessment and sorting of urban disaster vulnerability

Region	$A1$	$A2$	$A3$	$A4$	$B1$	$B2$	$C1$	$C2$	F	Rank
CDER	0.1656	0.4301	0.553	0.5751	0.4656	0.4836	0.3493	0.4725	0.6379	5
NESER	0.1917	0.5166	0.5923	0.6026	0.5081	0.6781	0.442	0.6049	0.6616	4
NWSER	0.5643	0.4379	0.3448	0.5118	0.5209	0.5456	0.6892	0.5062	0.8667	2
PXER	0.3514	0.4093	0.4762	0.5398	0.5687	0.4881	0.509	0.4058	0.7699	3
SSER	0.6636	0.6883	0.5596	0.5888	0.5209	0.5456	0.4379	0.5566	0.897	1

3.3 Evaluation Results and Comparative Analysis

The vulnerability distribution map, which is plotted according to the economic zones' vulnerability indexes, can visually help decision makers get regional vulnerability status and significant factors. Urban planning development and prevention of climate disasters can effectively reduce the city disaster vulnerability level. According to the calculation results of Table 2, our study draws the status of the disaster vulnerability, disaster response, sensitivity and resilience of each economic zone, as shown in Fig. 2.

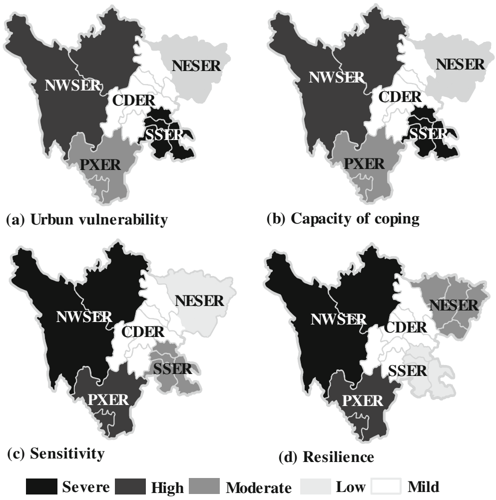


Fig. 2. Urban vulnerability of change-induced disasters

(1) Chengdu Economic Region (CDER)
CDER consists of 8 cities, including Chengdu, Deyang, Mianyang, Leshan, Meishan, Ziyang, Suining and Yaan. As shown in Fig. 1, the total GDP amount and urban built-up areas are more than half of the all urban corresponding indicators, the city's population of which is above one-third of

the total urban population. Therefore, the urban disaster vulnerability of CDER plays an important role in whole Sichuan urban climate disasters prevention systems.

Among all economic zones, the disaster vulnerability, coping ability, vulnerability, sensitivity and vulnerability resilience vulnerability indicators of the CDER all place in the lowliest place. The CDER's overall situation of climate disasters prevention is relatively benign. The value of the vulnerability index $A3$ and $A4$ is higher than other vulnerability indicators of CDER, which means disaster awareness and hospital conditions need to be improved. As the CDER is the core area of the entire Sichuan region, the overall level of medical care and economy is better than other economic regions. In the context of population mobility, the actual population of CDER is much larger than the data of the Yearbook, so the actual vulnerability is more serious than that in Table 2. And to minimize the urban vulnerability of the CDER, appropriate measures must be taken, especially on the disaster awareness and medical conditions. The urban management of CDER should offer much more attention to the floating population, and effectively improve the urban climate disaster vulnerability.

(2) Northeast Sichuan Economic Region (NESER)
The NESER includes 5 cities i.e., Guangyuan, Nanchong, Guangan, Dazhou and Bazhong. As shown in Fig. 2, the rank of the NESER's resilience vulnerability is more serious than that of coping vulnerability and sensitivity vulnerability. Disaster sensitivity of the NESER needs to be promoted, so that more energy should be put into the economic construction and the social security expenditure. Among them, the urban economic development is the most important link.

NESER is located in the northeast of Sichuan, which is an important channel to link Chengdu-Chongqing Economic Zone. NESER has rich mineral resources, especially natural gas reserves. The overall economic development level of ERSER is relatively backward; the communication and cooperation between cities in the region is poor, and at present there are no strong core cities to promote the economic development.

Relying on its geographical advantage and the construction of Chengdu-Chongqing Economic Zone, NESER can build wide-ranging and multi-level pattern of opening to the outside regions. The establishment of regional economic center will give priority to develop the core city, and injects impetus for economic development of the urban. These measures can not only promote the economic development of cities, but also can improve the disaster recovery, reduce the vulnerability of urban disasters, and increase the ability to withstand climate disasters.

(3) Northwest Sichuan Economic Region (NWSER)
NWSER is located in the northwest of Sichuan Province, of which the climate is plateau-climate. And the tourism resources and the natural resources are very rich. As a national autonomous region, the Northwest Economic

Zone covers a larger area than other economic zones. However, its urbanization process is slow. The ratio of the GDP, urban built-up area and urban population is much less than other areas (Fig. 1). As shown in Fig. 2, the sensitivity and vulnerability of the urban economic zone is very serious. The urban situation, the structure of population, self rescue ability and social security is poor.

In the vulnerability system, the population structure consists of higher education ratio, elderly population ratio and population density. And these indicators have a great relationship with local talent attractiveness, and the ratio of the young. In order to reduce the vulnerability, urban management departments need to increase investment in municipal construction, and improve urban facilities. On another hand, effective measures should be taken to make the urban appeal to the highly educated young and adults.

NWSER is an important tourist area in China, and its economy is mainly from agriculture sector. In view of these regional characteristics, the commercial organizations can build or upgrade related tourist facilities (e.g., the road, the gas stations and other infrastructure or public services), and make full use of these resources, so as to retain high-quality personnel and construct the ecological or tourist city.

(4) Pan-Xi Economic Region (PXER)
PXER is composed of Panzhihua and Liangshan. This economic region has a huge potential for economic development. As shown in Fig. 2, the overall urban vulnerability level of PXER is moderate in the five major economic zones. The root reasons of PXER's high vulnerability are similar to the causes of NWSER i.e., the less attractiveness and the shortage of relevant financial expenditure. And relative measures must be based on the improvement of the urban regional economic level.

In the background of that the central government of China strengthens the research of the vanadium and hydropower resources, this economic zone can take full of the mood of the time and transform the mainly strategic shift from the steel to the vanadium titanium new material with its abundant natural resources. To provide funding for municipal construction and enhance the attractiveness, the management of PXER should innovate the development model and increase revenue.

(5) South Sichuan Economic Region (SSER)
SSER is located in the upper reaches of the upper reaches of the Yangtze River, adjacent to Yunnan-Guizhou Plateau. The economic region is an important transportation hub, and has always been the most important materials trading center.

As shown in Fig. 2, the coping capacity vulnerability of SSER is the most serious among all economic zones, and need to be paid much more effort. The urban disaster response consists of grass-roots organizations, traffic conditions, disaster awareness and medical conditions, in which the importance of grass-roots organizations is higher than other indicators. In this assessment system, the primary organization is composed of enterprises and community committees. According to the data of Table 2, we can know that in order to reduce the urban vulnerability, the number of enterprises needs to be increased.

Although SSRS has a comprehensive transportation system, the level of development integration is low. And the overall advantages of this economic region can't be fully utilized in the long term. In addition, the transformation pace of the old industrial base is very slow.

By actively exploring new model of integration development, SSRS can slow the disaster vulnerability. The process of urban integration can be promoted by establishing the development fund of Southern Sichuan urban group. SSRS has a long wine culture. By supporting wine enterprises diversified development, intelligent manufacturing, upgrading of traditional manufacturing machinery and service oriented manufacturing transformation can be accelerated. Meanwhile, effective measures should be taken to develop high growth industry, and accelerate the exploration or development of shale gas. By various ways, enterprises can provide a good environment.

4 Conclusion

As an important part of disaster theory, vulnerability assessment is an important way to estimate the urban coping ability to disasters. To evaluate the urban vulnerability of climate change disasters, our study regards the five economic zones of Sichuan Province as objects and establishes the assessment systems.

Through data analysis and discussion, our study found that the economic zones should take appropriate measures based on its own characteristics to reduce the urban vulnerability. CDER needs to pay more attention to the floating population. By giving priority to the development of core cities, NESER can boost the whole urban economic development and enhance the disaster resilience. NWSER can make full use of its agriculture and tourism advantages, and provide much more support for enterprises engaged in deep processing of agricultural products or developing tourism resources. By the establishment of ecological cities, financial expenditure of NWSER can be guaranteed.

To slow the urban vulnerability, PXER needs to innovate the resource exploitation model, carry out the strategic transformation and increase the fiscal revenue. Relying on its rich history of wine culture, SSER can achieve the diversified development of wine enterprises. In addition, through creating a favorable environment for the development of enterprises, the disaster response capacity can be increased. And SSER can reduce the level of disaster vulnerability.

Acknowledgments. This research is supported by the Program of Social Science Foundation of Sichuan (Grant No. SC16C010). The authors are also grateful to Sichuan University for providing research funding (No. 2016SCU11036; No. skqy201640).

References

1. Bachelet D, Ferschweiler K et al (2016) Climate change effects on southern california deserts. J Arid Environ 127:17–29
2. Bellard C, Bertelsmeier C et al (2012) Impacts of climate change on the future of biodiversity. Ecol Lett 15(4):365–377
3. Bowen A, Cochrane S, Fankhauser S (2011) Climate change, adaptation and economic growth. Clim Change 113(2):95–106
4. Cutter SL (1996) Vulnerability to environmental hazards. Prog Hum Geogr 20(4):529–539
5. Füssel HM, Klein RJT (2006) Climate change vulnerability assessments: an evolution of conceptual thinking. Clim Change 75(3):301–329
6. Hallegatte S, Green C et al (2013) Future flood losses in major coastal cities. Nature Clim Change 3(9):802–806
7. Mazumdar J, Paul SK (2016) Socioeconomic and infrastructural vulnerability indices for cyclones in the Eastern Coastal States of India. Nat Hazards 82(3):1621–1643
8. She L, Wang G, Xu J (2013) Research of influential factors of urban vulnerability based on interpretative structural modeling: a case study of rescue and recovery in Japan earthquake. Areal Res Develop (in Chinese) 12:18
9. Shen G, Pimm SL et al (2015) Climate change challenges the current conservation strategy for the giant panda. Biol Conserv 190:43–50
10. Songhe X, Chuanfeng H, Lingpeng M (2015) Assessment and the key influencing factors analysis of urban disaster prevention system vulnerability. Soft Science 9:131–134 (in Chinese)
11. Thompson TM, Rausch S et al (2014) A systems approach to evaluating the air quality co-benefits of us carbon policies. Nature Clim Change 4(10):917–923
12. Wild M (2009) Global environmental change. J Geophys Res Atmos 114(11):733–743
13. Xu J, Xie H et al (2015) Post-seismic allocation of medical staff in the longmen shan fault area: case study of the Lushan earthquake. Environ Hazards 14(4):1–23
14. Yang G, Guo S et al (2016) Multi-objective operating rules for danjiangkou reservoir under climate change. Water Resour Manage 30(3):1183–1202
15. Yi L, Xi Z et al (2014) Analysis of social vulnerability to hazards in china. Environ Earth Sci 71(7):3109–3117
16. Yoon DK (2012) Assessment of social vulnerability to natural disasters: a comparative study. Nat Hazards 63(2):823–843
17. Zhang Q, Meng H (2014) Social vulnerability and poverty to climate change: a summary on foreign research. J China Agric Univ (in Chinese) 31:29–33

Judging Customer Satisfaction by Considering Fuzzy Random Time Windows in Vehicle Routing Problems

Yanfang Ma[1,2], Cuiying Feng[2], Jing Zhang[1,2], and Fang Yan[3(✉)]

[1] School of Economics and Management, Hebei University of Technology, Tianjing 300401, People's Republic of China
[2] College of Economics and Management, Zhejiang University of Technology, Hangzhou 310014, People's Republic of China
[3] School of Economic and Management, Chongqing Jiaotong University, Chongqing 400074, People's Republic of China
yanfang@cqjtu.edu.cn

Abstract. This article puts forward a membership function of the customer satisfaction based on fuzzy random time windows in vehicle routing problems. The objective is to confirm that all the customers are satisfied in an acceptable degree by judging the vehicle arriving time. More specifically, time windows given by customers are taken as fuzzy random variables in this paper. And then, fuzzy random theory has been used to describe customers' time windows. Finally, a measure function has been given to calculate customers' satisfaction based on fuzzy random time windows.

Keywords: Vehicle routing optimization · Customer satisfaction · Time windows · Fuzzy random variable

1 Introduction

Generally, in the classical VRP, there is a set of customers, each of whom have their own demands. Vehicles in the same condition at the depot deliver goods to these customers, and they are required to start and end at the depot. The objective of the classical VRP is to minimize the total cost by designing an optimal delivery route for each vehicle. Nowadays, the VRP is a common problem in almost every industry such as supply chain management and transport planning. With the development of modern technology, customer satisfaction has become a hot issue in the VRP.

In the classical VRP, delivery vehicles usually need to meet the following conditions: (1) Serve all customers using a minimum of vehicles; (2) Each customer is served by only one vehicle once; (3) Each vehicle starts and ends at the depot; (4)Total customer demand on each route cannot exceed the load capacity of the vehicle. There are several VRP variants, such as the multi-depot VRP [1,9], the

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_16

Pickup and Delivery VRP [10], and the VRP with Backhauls [16] and so on. In addition, some problems have pre-set time constraints on the period of the day in which deliveries should take place. These are known as Vehicle Routing Problems with Time Windows (VRPTW) which is a well known variant of the VRP.

The VRPTW is classified into two types, namely a VRP with a hard time window (VRPHTW) in which the time windows are hard constraints and where a route is not feasible if the service for any customer does not start within the limits established by the time window [3], and a VRP with a soft time window (VRPSTW), which is a relaxation of the VRPHTW, where the delivery of goods is allowed outside the time windows if a penalty is paid [4,5,12]. There are many uncertainties in real-life applications where hard time windows can be violated, while the "soft time window" can deal with these situations [13,14]. The travel time on each road section cannot be determined because of many uncertainties, such as traffic incidents, vehicle breakdowns, work zones, special events, and driver skills and experience [6].

In the past, the research used to consider these uncertainties random variables [6,15], but recently, researchers have been increasingly applying fuzzy membership functions to characterize the service level issues associated with time window violation in a vehicle routing problem, called the VRP with fuzzy time windows (VRPFTW) [2,7,14]. This research has highlighted the fact that fuzzy factors and stochastic factors can exist at the same time in the VRPTM. In this paper, we propose to use fuzzy random theory to describe the VRPTW, namely VRPFRTM.

This paper is organized as follows. In Sect. 2, we describe the problem which includes fuzzy random time windows. And then, in Sect. 3, we discuss the methods for dealing with these uncertainties, and propose a procedure to handle fuzzy random time windows. Next, we proposed the membership function of the customer satisfaction based on the fuzzy random time windows in Sect. 4. Finally, some concluding remarks are outlined in Sect. 5.

2 Fuzzy Random Time Windows

The time windows in the traditional VRPTW can be described as follows. Each customer needs to be served in a certain time window, and no delay is permitted. This certain time window can be described as $[e, l]$, where e and l are the earliest and latest time which the customer can accept, respectively. As to vehicle routing problem with soft time windows (VRPSTW), the vehicle can provide service before e and also after l. Some researchers in this field believe that the violation adds some costs and these costs are related to the degree of the violation [14].

Adopting soft time windows often means loss in customer satisfaction. Soft time windows certainly can bring some economic benefits, while dissatisfaction causes customer defection in a long term, which leads to economic loss. It is necessary to maintain the customer satisfaction on a certain level and minimize the cost in the mean time. Some researchers have use fuzzy theory to describe the customer satisfaction level, such as [14]. In their research, it is believed that the customer satisfaction level is related to the start time. If a customer requires that the service should be started in $[e, l]$ in traditional VRP with hard time windows, earlier than e or later than l is unacceptable, and the satisfaction level is 0; otherwise, the customer is satisfied and the satisfaction level is 1. While considering the soft time windows, the service starts earlier than e or later than l is permitted in some extent, but no earlier than EET or later than ELT [14]. Hence, if service starts between e and l, customer satisfaction level is 1; if service starts between EET and e or between l and ELT customer satisfaction level is valued in $[0, 1]$; otherwise, customer satisfaction level is 0. The customer satisfaction level from hard to soft time windows can be seen in Fig. 1.

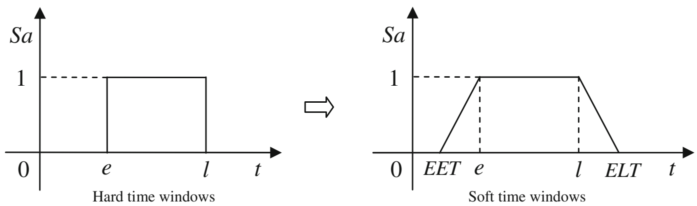


Fig. 1. The customer satisfaction level from hard to soft time windows

From previous studies, it can be seen that the EET and ELT are usually be treated as known and deterministic, while in practice, they can not usually be obtained as deterministic data. There are two common way to get the data, one is reasoning and the other is consulting the client. For reasoning, for example, the concreting in a construction project should be started at 9:00, and the process of unloading the concrete requires 10 min. By reasoning, the ELT is suppose to be 8:50, while actually the acceptable ELT to the project manager may be 8:30. As seen, reasoning often leads to false EET or ELT due to the lack of flexibility. As to consulting the client, deterministic data can not be obtained easily because the client usually gives a response including ambiguous information. The customer may make a statement, such as “not too early” or “no late than 10:00”. After getting these information, the specific data can not be acquired obviously, and if it is dealt as deterministic, the part of information which the customer provides will be lost. In sum, a response often not only includes some fuzzy information but also some random information. In this paper, the vehicle routing problem with fuzzy random time windows (VRPFRTM) is proposed and fuzzy random theory is chosen to describe the fuzzy random time windows. EET and ELT are considered to be fuzzy random, namely $\overline{\widetilde{EET}}$ and $\overline{\widetilde{ELT}}$.

3 Dealing with Fuzzy Random Variables

Uncertainty is absolute and certainty is relative. In practice, the fuzzy random variable has been applied in many areas. However, it is difficult to deal with uncertain variables, especially fuzzy random variables. Based on previous studies about fuzzy random theories, a method to transform a fuzzy random variable into a deterministic one is proposed in this paper. Take fuzzy random variable $\widetilde{\overline{EET}}$ as an example. The procedure of the transform operation can be described as follows:

Step 1. Consider the endurable earliness time of customer, $\widetilde{\overline{EET}}$, is fuzzy random variable. Through previous data and professional experience using statistical methods, estimate the parameters $[m]_L$, $[m]_R$, μ_0 and σ_0^2.

Step 2. Obtain the intermediate parameters by using group decision making approach, namely the decision-makers' degree of optimism. From Puri and Ralescu's definition in [11], a fuzzy random variable is a measurable function from a probability space to a collection of fuzzy variables. Roughly speaking, a fuzzy random variable is a random variable taking fuzzy values. In this paper, the fuzzy random variable $\widetilde{\overline{EET}}$ is denoted as $\widetilde{\overline{EET}} = ([m]_L, \rho(\omega), [m]_R)$, where $\rho(\omega) \sim N(\mu_0, \sigma_0^2)$ whose probability density function is $\varphi_\rho(x)$. Thus the expression of $\varphi_\rho(x)$ should be $\varphi_\rho(x) = \frac{1}{\sqrt{2\pi}\sigma_0 x} e^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}}$. Suppose that σ is a probability level and $\sigma \in [0, \sup \varphi_\rho(x)]$, r is a possibility variable and $r \in [r_l, 1]$ where $r_l = \frac{[m]_R - [m]_L}{[m]_R - [m]_L + \rho_\sigma^R - \rho_\sigma^L}$, both of them reflect the decision-maker's degree of optimism.

Step 3. Let ρ_σ be the σ-cut of the random variable $\rho(\omega)$. According to Xu and Liu's lemma in [17], $\rho_\sigma = [\rho_\sigma^L, \rho_\sigma^R] = \{x \in R | \varphi_\rho(x) \geq \sigma\}$, and the value of ρ_σ^L and ρ_σ^R can be expressed as

$$\rho_\sigma^L = \inf\{x \in R | \varphi_\rho(x) \geq \sigma\} = \inf \varphi_\rho^{-1}(\sigma) = \mu_0 - \sqrt{-2\sigma_0^2 \ln(\sqrt{2\pi}\sigma_0\sigma)},$$

$$\rho_\sigma^R = \sup\{x \in R | \varphi_\rho(x) \geq \sigma\} = \sup \varphi_\rho^{-1}(\sigma) = \mu_0 + \sqrt{-2\sigma_0^2 \ln(\sqrt{2\pi}\sigma_0\sigma)}.$$

Step 4. Transform the fuzzy random variables $\widetilde{\overline{EET}} = ([m]_L, \rho(\omega), [m]_R)$ into the (r, σ)-level trapezoidal fuzzy variable $\widetilde{\omega}_{\widetilde{\overline{EET}}(r,\sigma)}$ by the following equation:

$$\widetilde{\overline{EET}} \to \widetilde{\omega}_{\widetilde{\overline{EET}}(r,\sigma)} = ([m]_L, \underline{m}, \overline{m}, [m]_R).$$

here, we have

$$\underline{m} = [m]_R - r([m]_R - \rho_\sigma^L) = [m]_R - r([m]_R - \mu_0 + \sqrt{-2\sigma_0^2 \ln(\sqrt{2\pi}\sigma_0\sigma)}),$$

$$\overline{m} = [m]_R - r([m]_L + \rho_\sigma^R) = [m]_L + r(\mu_0 - [m]_L + \sqrt{-2\sigma_0^2 \ln(\sqrt{2\pi}\sigma_0\sigma)}).$$

$\widetilde{\overline{EET}}$ can be specified by $\tilde{\omega}_{\widetilde{\overline{EET}}(r,\sigma)} = ([m]_L, \underline{m}, \overline{m}, [m]_R)$ with the membership function:

$$\mu_{\tilde{\omega}_{\widetilde{\overline{EET}}(x)}} = \begin{cases} 0 & \text{for } \; x > [m]_R \\ \frac{[m]_R - x}{[m]_R - \overline{m}} & \text{for } \; \overline{m} \le x \le [m]_R \\ 1 & \text{for } \; \underline{m} \le x \le \overline{m} \\ \frac{x - [m]_L}{\underline{m} - [m]_L} & \text{for } \; [m]_L \le x \le \underline{m} \\ 0 & \text{for } \; x < [m]_L. \end{cases}$$

The process of transforming fuzzy random variable $\widetilde{\overline{EET}}$ to the (r, σ)-level trapezoidal fuzzy variable $\tilde{\omega}_{\widetilde{\overline{EET}}(r,\sigma)}$ is illustrated in Fig. 2.

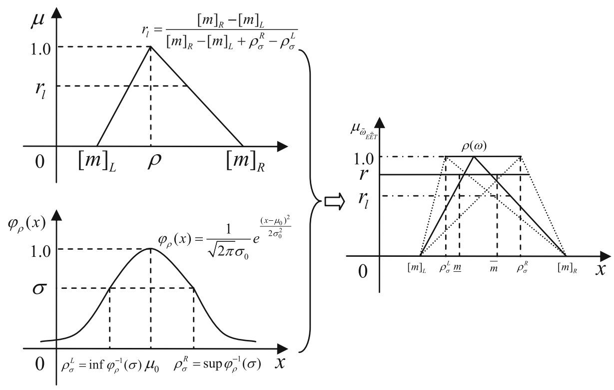


Fig. 2. The process of transforming fuzzy random variable

Step 5. Take expected value operator to convert (r, σ)-level trapezoidal fuzzy variable into a deterministic one. Based on the definition of fuzzy interval and expected value by Heilpern [8], suppose that there is a fuzzy number $\tilde{N} = (a, b, c, d)$, its membership function is:

$$\mu_{\tilde{N}}(x) = \begin{cases} 0 & \text{for } \; x < a \\ f_{\tilde{N}}(x) & \text{for } \; a \le x \le b \\ 1 & \text{for } \; b \le x \le c \\ g_{\tilde{N}}(x) & \text{for } \; c \le x \le d \\ 0 & \text{for } \; x > d. \end{cases}$$

where $f_{\tilde{N}}(x)$ and $g_{\tilde{N}}(x)$ are the upper and lower ends of the fuzzy number respectively. Then, the expected value of fuzzy variables $\tilde{\omega}_{\widetilde{\overline{EET}}(r,\sigma)}$ is as follows:

$$EV[\tilde{\omega}_{\widetilde{\overline{EET}}(r,\sigma)}] = \frac{1}{2}[(\underline{m} - \int_{[m]_L}^{\underline{m}} f_{\tilde{\omega}_{\widetilde{\overline{EET}}(r,\sigma)}}(x)dx) + (\overline{m} + \int_{\overline{m}}^{[m]_R} g_{\tilde{\omega}_{\widetilde{\overline{EET}}(r,\sigma)}}(x)dx)].$$

Based on the what has been discussed above, the fuzzy random time windows can be achieved, see Fig. 3.

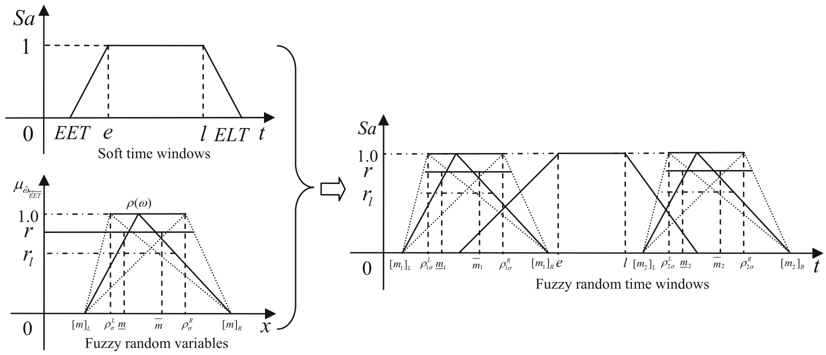


Fig. 3. The customer satisfaction level from soft to fuzzy random time windows

4 Customer Satisfaction by Fuzzy Random Time Windows

Finally, it comes how to define the customer satisfaction based on the fuzzy random time windows. In this paper, we suppose that there exists a least average customer satisfaction level Sa which is accepted by the customers [14].

$$\frac{1}{n}\sum_{i=1}^{n} L_i(t_i) \geq Sa, \tag{1}$$

where n is the number of customers. $L_i(t_i)$ is the satisfaction level serving customer i. Sa is the least average customer satisfaction level. The membership function of the satisfaction level, $L_i(t_i)$, is as below [14]:

$$L_i(t_i) = \begin{cases} 0, & t_i < \overline{\widetilde{EET}}_i, \\ \frac{t_i - \overline{\widetilde{EET}}_i}{e_i - \overline{\widetilde{EET}}_i}, & \overline{\widetilde{EET}}_i \leq t_i < e_i, \\ 1, & e \leq t_i < t, \\ \frac{\overline{\widetilde{ELT}}_i - t_i}{\overline{\widetilde{ELT}}_i - l_i}, & l_i \leq t_i < \overline{\widetilde{ELT}}_i, \\ 0, & t_i \geq \overline{\widetilde{ELT}}_i. \end{cases}$$

As discussed above, the service may start outside the time window $[e, l]$, and the bounds of acceptable earliness and lateness are described by $\overline{\widetilde{EET}}$ and $\overline{\widetilde{ELT}}$, respectively. Obviously, the earliness and lateness are closely related to the quality of service of the supplier. The response of a customer satisfaction level to a given service time may not be simply "good" or "bad"; instead, it may be between "good" and "bad". For example, the customer might say, "it's all right" to be served within $[\widetilde{EET}, e]$ or $[l, \widetilde{ELT}]$. In either case, the service level cannot be described by only two states (0 or 1).

5 Conclusion

In this paper, we focus on a special type of the VRPTW, the vehicle routing problem with fuzzy random time windows and multiple decision makers (VRPFRTW-MDM) which is seldom considered before. We present a membership function of the customer satisfaction based on fuzzy random time windows in vehicle routing problems. Since customer satisfaction is becoming more and more important for suppliers, the objective of this paper is to confirm that all the customers are satisfied in an acceptable degree by judging the vehicle arriving time. We also proposed a method to deal with the fuzzy random time windows based on the fuzzy random theory. In the end, we have given a measure function on how to obtain the customers' satisfaction based on fuzzy random time windows.

Acknowledgement. This research was supported by National Natural Science Foundation of China (Grant No. 71640013, and Grant No. 71401020), and System Science and Enterprise Development Research Center (Grant No. Xq16C10), National philosophy and Social Science Foundation of Tianjin (TJGL15-007); Provincial Science and Technology Foundation of Hebei Province (16457643D).

References

1. Aras N, Aksen D, Tuğrul Tekin M (2011) Selective multi-depot vehicle routing problem with pricing. Transp Res Part C Emerg Technol 19(5):866–884
2. Brito J, Martínez F et al (2012) Aco-grasp-vns metaheuristic for vrp with fuzzy windows time constraints. In: Computer aided systems theory–EUROCAST 2011, Springer, pp 440–447
3. Cordeau JF, Desaulniers G et al (2002) Vrp with time windows. Veh Routi Prob 9:157–193
4. Errico F, Desaulniers, et al (2016) The vehicle routing problem with hard time windows and stochastic service times. Euro J Trans Logist:1–29
5. Figliozzi MA (2010) An iterative route construction and improvement algorithm for the vehicle routing problem with soft time windows. Trans Res Part C Emerg Technol 18(5):668–679
6. Gao S, Chabini I (2006) Optimal routing policy problems in stochastic time-dependent networks. Trans Res Part B Methodol 40(2):93–122
7. Ghoseiri K, Ghannadpour SF, Seifi A (2010) Locomotive routing and scheduling problem with fuzzy time windows. In: Transportation research board 89th Annual Meeting (No: 10-2592)
8. Heilpern S (1992) The expected value of a fuzzy number. Fuzzy Sets Syst 47:81–86
9. Lightnerlaws C, Agrawal V et al (2016) An evolutionary algorithm approach for the constrained multi-depot vehicle routing problem. Int J Intell Comput Cybern 9(1)

10. Nair DJ, Grzybowska H, et al (2016) Food rescue and delivery a heuristic algorithm for periodic unpaired pickup and delivery vehicle routing problem. Trans Res Board
11. Puri ML, Ralescu DA (1986) Fuzzy random variables. J Math Anal Appl 114(2):409–422
12. Qureshi A, Taniguchi E, Yamada T (2009) An exact solution approach for vehicle routing and scheduling problems with soft time windows. Transp Res Part E Logist Transp Rev 45(6):960–977
13. Shi C, Li T et al (2016) A heuristics-based parthenogenetic algorithm for the vrp with potential demands and time windows. Sci Program 1:1–12
14. Tang J, Pan Z et al (2009) Vehicle routing problem with fuzzy time windows. Fuzzy Sets Syst 160(5):683–695
15. Taş D, Dellaert N et al (2012) Vehicle routing problem with stochastic travel times including soft time windows and service costs. Comput Oper Res
16. Wassan N, Wassan N et al (2016) The multiple trip vehicle routing problem with backhauls: formulation and a two-level variable neighbourhood search. Comput Oper Res
17. Xu J, Liu Y (2008) Multi-objective decision making model under fuzzy random environment and its application to inventory problems. Inform Sci 178(14):2899–2914

Hybrid Multiobjective Evolutionary Algorithm with Differential Evolution for Process Planning and Scheduling Problem

Chunxiao Wang[1,2], Wenqiang Zhang[1,2(✉)], Le Xiao[1,2], and Mitsuo Gen[3]

[1] Henan University of Technology, Zhengzhou, People's Republic of China
[2] Key Laboratory of Grain Information Processing and Control, Ministry of Education, Zhengzhou, People's Republic of China
zhangwq@haut.edu.cn
[3] Fuzzy Logic Systems Institute and Tokyo University of Science, Tokyo, Japan

Abstract. In an intelligent manufacturing environment, process planning and scheduling (PPS) plays a very important role as a most complex and practical scheduling problem, which processes a set of prismatic parts into completed products by determining the optimal process plans and moments to execute each operation with competitive manufacturing resources. Many research works use multiobjective evolutionary algorithm (MOEA) to solve PPS problems with consideration of multiple complicated objectives to be optimized. This paper proposes a hybrid multiobjective evolutionary algorithm with differential evolution (HMOEA-DE) to solve the PPS problem. HMOEA-DE uses a special designed fitness function to evaluate the dominated and nondominated individuals and divides the population and elitism into two parts, which close to the center and edge areas of Pareto frontier. Moreover, differential evolution applied on elitism tries to improve the convergence and distribution performances much more by guiding the search directions through different individuals with different fitness function. Numerical comparisons indicate that the efficacy of HMOEA-DE outperforms the traditional HMOEA without DE in convergence and distribution performances.

Keywords: Process planning and scheduling · Multiobjective evolutionary algorithm · Differential evolution

1 Introduction

In an intelligent manufacturing system, a set of prismatic parts need to be processed into a products effectively and economically according to various resources constraints. The parts have operations with different features, which to be related with the machines, tools and tool access directions (TADs). Moreover, the precedence relationship constraints among operations must be satisfied with the technological and the geometrical considerations. Process planning

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_17

generates the optimal process plans, i.e. optimizes the operations (machine, tool, tool access direction) and their sequences. The scheduling assignments the most appropriate moment to execute each operation with competitive resources. The process planning and scheduling (PPS) optimizes of the optimal process plan and schedule simultaneously within the precedence relationship constraints and manufacturing resources. However, it is very difficult to efficiently find an optimal solutions from all of the combinations of all of the operations, manufacturing resources satisfying one and more specified objectives.

Since the complex and multi-resource constraints as well as multi-objective requirements of PPS problem, many researchers have addressed the evolutionary algorithms (EAs) method to deal with it while satisfying the single and multiple objectives [3,10,11,17,19]. Especially, multi-objective evolutionary algorithms (MOEAs) are well suitable for solving multi-objective optimizations (MOPs) [6,18]. In MOEAs, Vector evaluated genetic algorithm (VEGA) divides the population into several sub-populations according to the number of objectives, each of which evolves toward a single objective [14]. As two classical MOEAs, NSGA-II [5] and SPEA2 [20] have been proven to get better convergence and distribution performances in solving MOPs. NSGA-II uses the Pareto ranking and crowding distance mechanism to obtain better performances. SPEA2 proposes raw fitness assignment mechanism and density mechanism to guarantee the convergence and distribution performances. For solving the multi-objective PPS problem, Zhang et al., propose a hybrid sampling strategy-based multi-objective evolutionary algorithm (HSS-MOEA), which combines the sampling strategy of VEGA and the sampling strategy based on a new Pareto dominating and dominated relationship-based fitness function (PDDR-FF) [19]. The hybrid sampling strategies preserve both the convergence and the distribution performance while reduce the computational time.

As a real-parameter optimization method, Differential Evolution (DE) [16] has been also treated as a population-based approach, in which mutation and crossover are the variation operators used to generate new solutions. In DE, the mutation operator has been executed based on the differences of individuals to guide the search direction in the current population and the replacement mechanism uses to maintain the population. The effectiveness but simplicity of DE has attracted many research interests in single objective and multi-objective optimization [1,2,4,7–9,12,13,15]. Especially, Iorio and Li [8] proposed three DE variants incorporating the directional information to improve the optimization performance. The direction information among different individuals with same or different rank as well as different crowding distances could guide the search converge towards the Pareto frontier and/or prefers along the Pareto frontier to improve the convergence and distribution performances. Iorio and Li [8] provided an interesting topic to combine the DE with other MOEAs to improve the performance more especially for most complicated PPS problem, since only applies the DE on PPS problem could increase the computation time for practical optimization problem.

This paper proposes a hybrid multi-objective evolutionary algorithm with differential evolution (HMOEA-DE) to solve the multi-objective PPS problem that combines the DE with HSS-MOEA. The HSS-MOEA could generate the new individuals with traditional crossover and mutation operators, and update the elitism with special fitness function and strategy. The DE could enhance the performance on elitist population after applying the HSS-MOEA.

The paper is organized as follows: Sect. 2 presents the PPS problem and formulates the mathematical model; Sect. 3 describes the detailed HMOEA-DE approach; Sect. 4 gives a discussion and analysis of numerical experiments results; finally, the conclusion are drawn in Sect. 5.

2 Process Planning and Scheduling Problem

In PPS, there are many parts to be processed by a number of available machines with different tools and different TADs. Each part has several operations, and each operation can be processed on machines with different processing times. The PPS problem finally generates a proper process plan and schedule, which tells decision maker how, when and in which sequence to allocate suitable manufacturing resources to operations effectively.

The operation sequence of PPS should be decided firstly, then, the manufacturing resources should be assigned to each operation in operation sequence. After generating the process plan, the PPS determine the most appropriate moment to process each operation with competitive resources. The precedence relationship and resource constraints must be satisfied in deciding the process plan and schedule. Moreover, the process plan and schedule should also satisfy the multiple objectives requirements to maintain the feasibility.

Two important and frequently-used objectives are minimizing makespan and variation of workload for each machine.

The mathematical model of the problem is expressed in the following notations.
Indices:

i, k: indices of part, $(i, k = 1, 2, \cdots, I)$.
j, h: indices of operation for part, $(j, h = 1, 2, ..., J_i)$.
m: index of machine, $(m = 1, 2, \cdots, M)$;
l: index of tool, $(l = 1, 2, \cdots, L)$.
d: index of TAD, $(d = 1, 2, \cdots, D)$.

Parameters:

I: number of parts;
J_i: number of operations for part i;
M: number of machines;
L: number of tools;
D: number of TADs;
O: set of operations for part i, $O_i = \{o_{ij} | j = 1, 2, \cdots, J_i\}$;

o_{ij}: the j$^{\text{th}}$ operation of part;
m_m: the m$^{\text{th}}$ machine;
t_l: the l$^{\text{th}}$ tool;
a_d: the d$^{\text{th}}$ TAD;
M_{ij}: set of machines that can process o_{ij};
A_m: set of operations that can be processed on machine m;
S: operation sequence. $S = (s_1, s_2, \cdots, s_w, \cdots, s_W), s_w = o_{ij}, W = \sum_{i=1}^{I} J_i$;
r_{ijh}: precedence constraints. If o_{ij} is predecessor of o_{ih}, $r_{ijh} = 1$; otherwise, 0;
t_{mij}^{M}: machining time of $o^i j$ by machine m;
t^{MCI}: machine change time index. It is the same for each machine change.
t_{mij}^{MC}: machine change time of o_{ij} by machine m. A machine change is described in article by Zhang et al. [19]. Once machine change is occurred, $t_{mij}^{\text{MC}} = t^{\text{MCI}}$; otherwise, $t_{mij}^{\text{MC}} = 0$;
t^{TCI}: tool change time index. It is the same for each tool change;
t_{mij}^{TC}: tool change time of o_{ij} by machine m. A tool change is described in article by Zhang et al. [19] If tool change is occurred, $t_{mij}^{\text{TC}} = t^{\text{TCI}}$; otherwise, $t_{mij}^{\text{TC}} = 0$;
t^{SCI}: set-up change time index. It is the same for each set-up change;
t_{mij}^{SC}: set-up change time of o_{ij} by machine m. A set-up change is described in article by Zhang et al. [19] If set-up change is occurred, $t_{mij}^{\text{SC}} = t^{\text{SCI}}$; otherwise, $t_{mij}^{\text{SC}} = 0$;
t_{mij}^{PRE}: preparation time of operation o_{ij} by machine m. The preparation time for an operation consists of machine change time, tool change time and set-up time for the operation;
t_{mij}^{P}: processing time of operation o_{ij} by machine m.

$$t_{mij}^{\text{PRE}} = t_{mij}^{\text{MC}} + t_{mij}^{\text{TC}} + t_{mij}^{\text{SC}} \tag{1}$$

The processing time for an operation consists of the preparation time and the machining time for the operation.

$$t_{mij}^{\text{P}} = t_{mij}^{\text{PRE}} + t_{mij}^{\text{M}}, \tag{2}$$

where u_m is workload of machine m.

$$u_m = \sum\nolimits_{i=1}^{I} \sum\nolimits_{j=1}^{J_i} t_{mij}^{\text{M}} x_{mij}^{\text{M}}, \tag{3}$$

where $\bar{u}$: average workload of machine.

$$\bar{u} = {}^1\!/_M \sum\nolimits_{m=1}^{M} u_m, \tag{4}$$

where t_{mij}^{C}: completion time of $o_i j$ by machine m, it should satisfy the inequality $t_{m'i(j-1)}^{\text{C}} + t_{mij}^{\text{P}} \leq t_{mij}^{\text{C}}$ that means for every operation, its direct predecessor's completion time plus its processing time might be smaller that its completion time.

Decision variables:

$x_{mij}^{M} = \begin{cases} 1, \text{ if } o_{ij} \text{ is performed by machine} m, \\ 0, \text{ otherwise;} \end{cases}$

$x_{lij}^{T} = \begin{cases} 1, \text{ if } o_{ij} \text{ is performed by tool } l, \\ 0, \text{ otherwise;} \end{cases}$

$x_{dij}^{D} = \begin{cases} 1, \text{ if } o_{ij} \text{ is performed by TAD } d, \\ 0, \text{ otherwise;} \end{cases}$

$y_{ijkh} = \begin{cases} 1, \text{ if } o_{ij} \text{ is performed directly before } o_{kh}, \\ 0, \text{ otherwise;} \end{cases}$

$\Omega\left(X, Y\right) = \begin{cases} 1, \text{ if } X \neq Y, \\ 0, \text{ otherwise.} \end{cases}$

The mathematical model can be formulated as the following bi-criteria nonlinear mixed integer programming (NMIP) model:

$$\min \ t_M = \max_{m,i,j}\{t_{mij}^{C}\} \tag{5}$$

$$\min \ w_P = \sqrt{1/M \sum\nolimits_{m=1}^{M} \left(u_m - \bar{u}\right)^2} \tag{6}$$

$$\text{s. t.} \ \left(t_{mkh}^{\mathrm{C}} - t_{mkh}^{\mathrm{P}} - t_{mij}^{\mathrm{C}}\right) x_{mij}^{M} x_{mkh}^{M} y_{ijkh} \geq 0, \ \forall\left(i,j\right),\left(k,h\right),m \tag{7}$$

$$r_{ijh} y_{ihij} = 0, \ \forall(i,j), h \tag{8}$$

$$y_{ijij} = 0, \ \forall(i,j) \tag{9}$$

$$\sum\nolimits_{m=1}^{M} x_{mij}^{M} = 1, \ \forall(i,j) \tag{10}$$

$$x_{mij}^{M} = 0, \ \forall\left(i,j\right) \notin A_m, \forall m \tag{11}$$

$$y_{ijkh} \in \{0,1\}, \ \forall\left(i,j\right),\left(k,h\right) \tag{12}$$

$$x_{mij}^{M} \in \{0,1\}, \ \forall m, \left(i,j\right) \tag{13}$$

$$t_{mij}^{\mathrm{C}} \geq 0, \ \forall m, \left(i,j\right). \tag{14}$$

Equation illustrates the minimization of makespan. Makespan is the maximization of completion time among all the operations. Minimization of workload variance is defined as Equation. Workload variance is defined as standard deviation of workload of all the machines. The equations to impose the precedence relationship constraints, resource constraints and nonnegative condition.

3 Hybrid Multi-Objective Evolutionary Algorithm with Differential Evolution

The solution procedure of HMOEA-DE includes 5 phases.

Phase 1: Selection of Better Individuals with Objective Values

The sampling strategy of VEGA is used to select the better individuals from population $P(t)$ into part of mating pool (sub populations) according to their objectives value.

Phase 2: Generation of Mating Pool with Hybrid Sampling

The sub populations and elitist population $A(t)$ are combined to form a mating pool. In the mating pool, sub-pop-1 saves the good individuals for one objective, and sub-population 2 stores the good individuals for the other objective. The elitist population (archive) holds the individuals with good PDDR-FF values. Therefore, one-third of the individuals serve one objective, one-third the other objective, and the left one-third both the two objectives in the mating pool.

Phase 3: Reproduction with Traditional Crossover and Mutation Operators

Problem-dependent crossover and mutation operators are used to reproduce new individuals to form new population P(t+1). Moreover, a local search mechanism is proposed to improve the quality of individuals after the reproduction process. The details of reproduction and local search is described in article by Zhang et al. [19].

Phase 4: Archive Maintenance by PDDR-FF based Elitist Sampling Strategy

The individuals in $A(t)$ and $P(t)$ are combined to form a temporary archive $A'(t)$. Thereafter, the PDDR-FF values of all individuals in $A'(t)$ are calculated and sorted. If the individual is nondominated one, its fitness value will not exceed one. The fitness value of dominated individual will exceed one. Moreover, even the all nondominated individuals have also different PDDR-FF values. The nondominated individuals will be smaller values (near to 0) than the edge points (near to 1). After calculating the fitness function value, the smallest $|A(t)|$ individuals in $A'(t)$ are copied to form $A'(t+1)$.

Phase 5: Archive Enhancement by Combining DE

In this paper, the DE operators such as mutation, crossover and selection operators only apply on the temporary archive A'(t+1), which like a local search process to improve the performance of archive. After DE operator, the obtained solution set (temporary archive) $A''(t+1)$ and $A'(t+1)$ are combined to generate the last $A(t+1)$ according to the PDDR-FF values.

The DE/current/2 is used to generate new mutants and each individuals x_i in $A'(t+1)$ and three individuals r_1, r_2, r_3 are randomly picked up to do reproduction operators according to the following mutation operator.

$$\mathbf{v}_i = \mathbf{x}_i + K(\mathbf{r}_3 - \mathbf{x}_i) + F(\mathbf{r}_1 - \mathbf{r}_2), \qquad (15)$$

where $x_i \neq r_1 \neq r_2 \neq r_3$, eval($r_1$)< eval($r_2$) and eval($r_3$)<eval($x_i$), K and F are scale factors and $0 \leq K,\ F \leq 1$. eval(x_i) means the fitness function value of x_i. It is important to note that in PDDR-FF, the smaller value mean better and the fitness function of nondominated individual will not exceed one. The fitness value of dominated individual will exceed one. Moreover, the nondominated individuals locating around the central region of Pareto frontier have smaller values (near to 0) than the edge points (near to 1).

The mutation operator of DE will improve the performances as follows schemes according to the different fitness function of current and three random individuals.

Scheme 1: the current individual x_i is nondominated solution (eval(s_i) ≤ 1).

(1) If the eval(r_1) ≤ 1 and eval(r_2) ≤ 1, r_1, r_2 are nondominated solutions. This mutation means the search direction starts from the nondominated solution (x_i) and firstly follows the direction from edge region to center region(x_i locates in edge region and r_3 locates in center region) along the Pareto frontier, next, the search direction will guide from edge region to center region (r_2 locates in edge region and r_1 locates in center region) again. It will improve the distribution performance.
(2) If the eval(r_1) > 1 and eval(r_2) > 1, r_1, r_2 are dominated solutions. The search direction starts from the nondominated solution (x_i) and firstly follows the direction from edge region to center region(x_i locates in edge region and r_3 locates in center region) along the Pareto frontier, next, the search direction will guide from edge region to center region (r_2 locates in edge region and r_1 locates in center region) again. It will improve the distribution performance.
(3) If the eval(r_1) ≤ 1 and eval(r_2) > 1, r_1 is nondominated solutions and r_2 is dominated solutions. The search direction starts from the nondominated solution (x_i) and firstly follows the direction from edge region to center region(x_i locates in edge region and r_3 locates in center region) along the Pareto frontier to improve the distribution performance, next, the search direction will guide from dominated region to nondominated region (r_2 locates in the dominated region and r_1 locates in the nondominated region) towards the Pareto frontier to improve the convergence performance.

Scheme 2: the current individual x_i is dominated solution (eval(s_i) > 1) and r_3 is nondominated solution (eval(r_3) < 1).

(1) If the eval(r_1) ≤ 1 and eval(r_2) ≤ 1, r_1, r_2 are nondominated solutions. This mutation means the search direction starts from the dominated solution (x_i) and firstly follows the direction from dominated region to nondominated region (x_i locates in the dominated region and r_3 locates in the nondominated region) along the Pareto frontier to improve the convergence performance, next, the search direction will guide from edge region to center region (r_2 locates in edge region and r_1 locates in center region) to improve the distribution performance.
(2) If the eval(r_1) > 1 and eval(r_2) > 1, r_1, r_2 are dominated solutions. The search direction starts from the dominated solution (x_i) and firstly follows the direction from dominated region to nondominated region (x_i locates in the dominated region and r_3 locates in the nondominated region) along the Pareto frontier to improve the convergence performance, next, the search direction will guide from edge region to center region (r_2 locates in edge region and r_1 locates in center region) to improve the distribution performance.
(3) If the eval(r_1) ≤ 1 and eval(r_2) > 1, r_1 is nondominated solutions and r_2 is dominated solutions. The search direction starts from the dominated solution (x_i) and firstly follows the direction from dominated region to nondominated region (x_i locates in the dominated region and r_3 locates in the

nondominated region) along the Pareto frontier to improve the convergence performance, next, the search direction will guide from dominated region to nondominated region (r_2 locates in the dominated region and r_1 locates in the nondominated region) towards the Pareto frontier to improve the convergence performance more.

Scheme 3: the current individual x_i is dominated solution ($\text{eval}(s_i) > 1$) and r_3 also is dominated solution ($\text{eval}(r_3) > 1$).

(1) If the $\text{eval}(r_1) \leq 1$ and $\text{eval}(r_2) \leq 1$, r_2 are nondominated solutions. This mutation means the search direction starts from the dominated solution (x_i) and firstly follows the direction from edge region to center region(x_i locates in edge region and r_3 locates in center region) along the Pareto frontier, next, the search direction will guide from edge region to center region (r_2 locates in edge region and r_1 locates in center region) again. It will improve the distribution performance.
(2) If the $\text{eval}(r_1) > 1$ and $\text{eval}(r_2) > 1$, r_1, r_2 are dominated solutions. The search direction starts from the dominated solution (x_i) and firstly follows the direction from edge region to center region(x_i locates in edge region and r_3 locates in center region) along the Pareto frontier, next, the search direction will guide from edge region to center region (r_2 locates in edge region and r_1 locates in center region) again. It will improve the distribution performance.
(3) If the $\text{eval}(r_1) \leq 1$ and $\text{eval}(r_2) > 1$, r_1 is nondominated solutions and r_2 is dominated solutions. The search direction starts from the dominated solution (x_i) and firstly follows the direction from edge region to center region(x_i locates in edge region and r_3 locates in center region) along the Pareto frontier to improve the distribution performance, next, the search direction will guide from dominated region to nondominated region (r_2 locates in the dominated region and r_1 locates in the nondominated region) towards the Pareto frontier to improve the convergence performance.

From above analysis, the HSS-MOEA could improve the convergence and distribution performance, the combination of HMOEA and DE could much improve the performances more.

4 Experiments and Discussion

In this paper, 4 parts (each part having 20, 16, 14 and 7 operations, respectively) problem are used. All the simulations are run on Core i5 processor (2.6 GHz clock). The adopted parameters are listed as follows: population size = 100; maximum generation = 500; archive size = 50; crossover probability = 0.70; mutation probability = 0.30 and the scale factor K and F are assigned randomly within 0 to 1. HMOEA-DE and HSS-MOEA are run 30 times to compare the results.

Let S_j be a solution set for each method ($j = 1, 2$). PF^* is a known reference Pareto set which comes from combining all of the obtained Pareto set by

2 methods with 30 runs. In this study, Coverage $C(S_1, S_2)$, Average distance $AD(S_j)$, Hyper volume $HV(S_j)$ and Spacing $SP(S_j)$ are selected to compare the performances of algorithms.

Figure 1 shows the numerical comparison of the box-and-whisker plots for C, AD, HV, and SP by HMOEA-DE, and HSS-MOEA. From Fig. 1, it is easy to see that the HMOEA-DE is better than HSS-MOEA in terms of C measure. The AD measure also indicates that HMOEA-DE can get smaller value than HSS-MOEA. For HV measure, HMOEA-DE is better than HSS-MOEA. The distribution performances, SP, indicates that HMOEA-DE is better than HSS-MOEA. The PDDR-FF could give different values among nondominated and dominated solution in elitism, and the combination of DE for individuals in elitism could significantly improve the convergence and distribution performances.

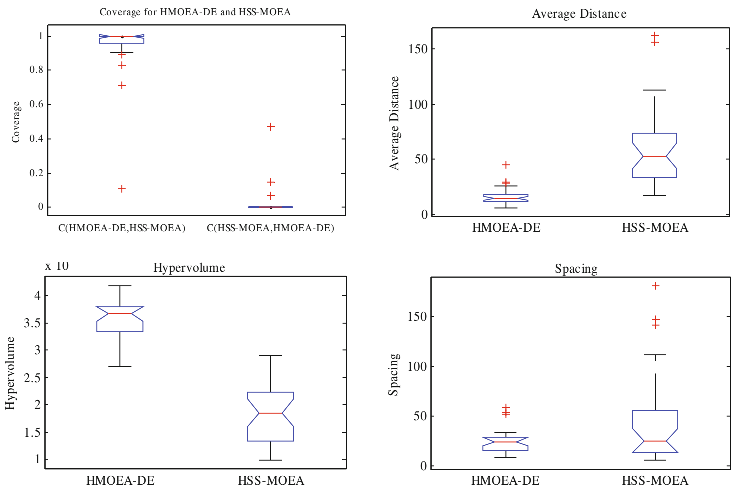


Fig. 1. C, AD, HV, and SP by HMOEA-DE, and HSS-MOEA

5 Conclusions

In this study, a HMOEA-DE approach was proposed to solve the multi-objective PPS problem with consideration of minimization of makespan and minimization of variation of machine workload simultaneously. The VEGA take a preference for the edge region of the Pareto front and the PDDR-FF-based elitism maintain mechanism has the tendency converging toward the center area of the Pareto front. Moreover, the hybrid DE operators for the elitism has the capabilities of improving the convergence and distribution performance much more. Numerical comparisons result demonstrated that HMOEA-DE outperformed the HSS-MOEA on convergence and distribution performance.

Acknowledgements. This research work is supported by the National Natural Science Foundation of China (U1304609), Foundation for Science & Technology Research Project of Henan Province (162102210044, 152102210068 152102110076), Program for Innovative Research Team (in Science and Technology) in University of Henan Province (17IRTSTHN011), Program for Key Project of Science and Technology in University of Education Department of Henan Province (17A520030), Fundamental Research Funds for the Henan Provincial Colleges and Universities (2014YWQQ12, 2015XTCX03, 2015XTCX04), Research Funds for Key Laboratory of Grain Information Processing and Control (Henan University of Technology) (KFJJ-2015-106), Ministry of Education and the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS): No. 15K00357.

References

1. Abbass HA, Sarker R (2011) The pareto differential evolution algorithm. Int J Artif Intell Tools 11(04):531–552
2. Ali M, Siarry P, Pant M (2012) An efficient differential evolution based algorithm for solving multi-objective optimization problems. Eur J Oper Res 217(2):404–416
3. Ausaf MF, Gao L, Li X (2015) Optimization of multi-objective integrated process planning and scheduling problem using a priority based optimization algorithm. Front Mech Eng 10(4):1–13
4. Babu BV, Jehan MML (2003) Differential evolution for multi-objective optimization. In: The 2003 Congress on Evolutionary Computation, 2003, CEC 2003, vol 4, pp 2696–2703
5. Deb K, Pratap A et al (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6(2):182–197
6. Gen M, Cheng R, Lin L (2008) Network models and optimization: Multiobjective genetic algorithm approach
7. Guo W, Yu X (2014) Non-dominated sorting differential evolution with improved directional convergence and spread for multiobjective optimization. In: Proceedings of the Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation, pp 87–88
8. Iorio AW, Li X (2006) Incorporating directional information within a differential evolution algorithm for multi-objective optimization. In: Conference on Genetic and Evolutionary Computation, pp 691–698
9. Li H, Zhang Q (2006) A multiobjective differential evolution based on decomposition for multiobjective optimization with variable linkages. In: Parallel Problem Solving From Nature - PPSN Ix, International Conference, Reykjavik, Iceland, 9–13 September 2006, Proceedings, pp 583–592
10. Nayak A (2015) Multi-objective process planning and scheduling using controlled elitist non-dominated sorting genetic algorithm. Int J Prod Res 53(6):1712–1735
11. Phanden RK, Jain A, Verma R (2013) An approach for integration of process planning and scheduling. Int J Comput Integr Manufact 26(4):284–302
12. Qu BY, Suganthan PN, Liang JJ (2012) Differential evolution with neighborhood mutation for multimodal optimization. IEEE Trans Evol Comput 16(5):601–614
13. Santana-Quintero LV, Coello Coello CA (2005) An algorithm based on differential evolution for multi-objective problems. Int J Comput Intell Res 1(2):151–169
14. Schaffer JD (1985) Multiple objective optimization with vector evaluated genetic algorithms. In: International Conference on Genetic Algorithms, pp 93–100

15. Sindhya K, Ruuska S et al (2011) A new hybrid mutation operator for multiobjective optimization with differential evolution. Soft Comput 15(10):2041–2055
16. Storn R, Price K (1997) Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 11(4):341–359
17. Xia H, Li X, Gao L (2016) A hybrid genetic algorithm with variable neighborhood search for dynamic integrated process planning and scheduling. Comput Ind Eng 102:99–112
18. Yu X (2010) Introduction to evolutionary algorithms. In: International Conference on Computers and Industrial Engineering, p 1
19. Zhang W, Gen M, Jo J (2014) Hybrid sampling strategy-based multiobjective evolutionary algorithm for process planning and scheduling problem. J Intell Manufact 25(5):881–897
20. Zitzler E, Laumanns M, Thiele L (2001) SPEA2: Improving the strength pareto evolutionary algorithm

A New Approach for Solving Optimal Control Problem by Using Orthogonal Function

Akram Kheirabadi[1](✉), Asadollah Mahmoudzadeh Vaziri[1], and Sohrab Effati[2]

[1] Department of Mathematics, Faculty of Mathematical Science and Statistics, University of Birjand, Birjand, Iran
ak_kheirabadi@yahoo.com

[2] Department of Applied Mathematics, Faculty of Mathematical Science and Statistics, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract. In the present paper we introduce a numerical technique for solving fractional optimal control problems (FOCP) based on an orthonormal wavelet. First we approximate the involved functions by Sine-Cosine wavelet basis; then, an operational matrix is used to transfer the given problem in to a linear system of algebraic equations. In fact operational matrix of the Riemann-Liouville fractional integration and derivative of Sine-Cosine wavelet are employed to achieve a linear algebraic equation, in place of the dynamical system in terms of the unknown coefficients. The solution of this system, gives us the solution of original problem. A numerical example is also given.

Keywords: Fractional optimal control problem · Sine-Cosine wavelet · Operational matrix · Caputo derivative · Riemann-Liouville fractional integration

1 Introduction

Many application of the fractional calculus is in basic sciences and engineering. Many realistic model of physical [8] phenomenon which has dependence at both the time instance and on the previous time history, can be utter with fractional calculus. For example it can be applied in nonlinear oscillations of earthquakes, fluid-dynamic traffic [9], frequency dependent damping behavior of various viscoelastic materials [2], solid mechanics [18], economics [3], signal processing [17], and control theory [4].

One of the main difficulties is how to solve the fractional differential equations. The most commonly techniques proposed to solve them are Adomian decomposition method (ADM) [22], Variational Iteration Method (VIM) [20], Operational Matrix Method [19], Homotopy Analysis Method [6,7], Fractional Difference Method (FDM) [15] and Power Series Method [16].

A fractional optimal control problem is an optimal control problem in which the performance index or the differential equations governing the dynamic of the

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_18

system or both contains at least one fractional order derivative term [25]. Integer order optimal controls have already been well established and a significant amount of works have been done in the field of optimal control of integer order systems. Agrawal formulated and developed a numerical scheme for the solution of FOCP [1] in the Caputo sense. Biswas proposed a pseudo-state space representation of a fractional dynamical system, which is exploited to solve a fractional optimal control problem using a direct numerical method [21]. Sweilam et al. solved some types of fractional optimal control problem with a Hamiltonian formula using a spectral method based on Chebyshev polynomials [24]. Bernstein polynomials have been used for finding the numerical solution of FOCP by using Lagrange multipliers [10].

Approximation by orthogonal families of basis functions is widely used in science and engineering. The main idea behind applying an orthogonal basis is reduction of the problem under consideration into a system of algebraic equations. This is possible by truncating series of orthogonal basis functions for the solution of the problem and applying operational matrices. The orthogonal functions are classified into three main category [23]: the first one is sets of piecewise constant orthogonal functions such as the Walsh functions and block pulse functions. The second one is orthogonal polynomials such as the Laguerre, Legendre and Chebyshev functions, and the last one is sine-cosine functions. In one hand approximating a continuous function with piecewise constant basis functions results in a piecewise constant approximation, on the other hand, if a discontinuous function is approximated with continuous basis functions, the resulting approximation is continuous which cannot properly model the discontinuities. So, neither continuous basis functions nor piecewise constant basis functions, if used alone, can efficiently model both continuity and discontinuity of phenomena at the same time. In the case that the function under approximation is not analytic, wavelet functions will be more effective.

In this paper, we propose a computational method based on Sine-Cosine wavelet with their fractional integration and derivative operational matrix to solve the FOCP. The main idea is reduction the problem under consideration into a system of algebraic equations. To this end, we expand the fractional derivative of the state variable and the control variable using the Sine-Cosine wavelet with unknown coefficients.

The paper is organized as follows. In first section we will give the definitions of fractional calculus, then express a brief review of block pulse function and the related fractional operational matrices. In Sect. 4, we describe Sine-Cosine wavelets and its application in function approximation. In Sect. 5, operational matrices of fractional integration and derivative for considered wavelet is given. In Sect. 6, the proposed method is described for solving the underlying FOCP. In the last section the proposed method is applied for solving numerical example.

2 Preliminaries of Fractional Calculus

The Riemann-Liouville fractional integration and Caputo differential operator of a function f of order $\alpha \geq 0$ is defined in [13] as:

$$(I^{\alpha}f)(t) = \begin{cases} \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau)d\tau & \alpha > 0 \\ f(t) & \alpha = 0, \end{cases} \tag{1}$$

$$\begin{aligned} D^{\alpha}f(t) &= \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-\tau)^{n-\alpha-1} f^{(n)}(\tau)d\tau \\ &= I^{n-\alpha}f^{n}(t) \quad n-1 < \alpha \leq n. \end{aligned} \tag{2}$$

3 Review of Block Pulse Functions and the Related Fractional Operational Matrix

In this section first we introduce block pulse function (BPF), then it's operational matrix of fractional integration.

3.1 Definition of BPF

A set of BPFs $B_{m'}(t)$ containing m' component functions in the interval [0, T) is given by

$$B_{m'}(t) \triangleq [b_0(t)b_1(t)\cdots b_i(t)\cdots b_{m'-1}]^T. \tag{3}$$

The ith component of the BPF vector $B_{m'}(t)$ is defined as

$$b_i(t) = \begin{cases} 1 & \frac{iT}{m'} \leq t < \frac{(i+1)T}{m'} \\ 0 & O.W. \end{cases} \quad i = 0, 1, 2, \cdots, m'-1. \tag{4}$$

A square integrable function f can be expanded into a BPF series as

$$f(t) = [c_0c_1\cdots c_i\cdots c_{m'-1}]B_{m'}(t) = C^TB_{m'}(t), \tag{5}$$

$$c_i = \frac{1}{h}\int_{ih}^{(i+1)h} f(t)dt \quad h = \frac{T}{m'}. \tag{6}$$

3.2 Operational Matrix for Fractional Integration of BPF

Suppose that F^{α} be the block pulse operational matrix of fractional integration [12]. It is defined as follows,

$$F_{\alpha} = h^{\alpha}\frac{1}{\Gamma(\alpha+2)}\begin{bmatrix} 1 & \xi_1 & \xi_2 & \cdots & \xi_{m'-1} \\ 0 & 1 & \xi_1 & \cdots & \xi_{m'-2} \\ 0 & 0 & 1 & \cdots & \xi_{m'-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \xi_1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7}$$

$$\xi_k = (k+1)^{\alpha+1} - 2k^{\alpha+1} + (k-1)^{\alpha+1} \quad k = 1, 2, \cdots, m'-1. \tag{8}$$

4 Description of Sine-Cosine Wavelets and Its Application in Function Approximation

4.1 The Sine-Cosine Wavelet

Sine-cosine wavelets $\psi_{n,m}(t)$ are defined as follows [11],

$$\psi_{n,m}(t) = \begin{cases} 2^{\frac{k+1}{2}} f(2^k t - n) & \frac{n}{2^k} \le t \le \frac{n+1}{2^k} \\ 0 & o.w \end{cases} \tag{9}$$

with

$$f_m(t) = \begin{cases} \frac{1}{\sqrt{2}} & m = 0 \\ \cos(2m\pi t) & m = 1, 2, \cdots, l \\ \sin(2(m-l)\pi t) & m = l+1, \cdots, 2l \end{cases} \tag{10}$$

$n = 0, 1, \cdots, 2^k - 1, k = 0, 1, \cdots,$ where l is any positive integer.

4.2 Function Approximation

A function $f(t) \in L^2[0, 1)$ can be approximated as:

$$f(t) = \sum_{n=0}^{2^k-1} \sum_{m=0}^{2l} c_{n,m}\psi_{n,m} = C^T \Psi(t) = \Psi^T(t)C, \tag{11}$$

where $c_{n,m} = \langle f(t), \psi_{n,m}\rangle$ and $\langle ., .\rangle$ denotes the inner product as:

$$c_{n,m} = \int_{-\infty}^{+\infty} f(t)\psi_{n,m}(t)dt. \tag{12}$$

where $\Psi(t)$ represent considered wavelet. C and $\Psi(t)$ are $2^k(2l+1) \times 1$ matrices which are given by:

$$C^T = [c_{00}c_{01} \cdots c_{0,2l}, c_{10}, \cdots, c_{1,2l}, \cdots, c_{2^k-1,0}, \cdots, c_{2^k-1,2l}], \tag{13}$$

$$\Psi^T = [\psi_{00}\psi_{01} \cdots \psi_{0,2l}, \psi_{10}, \cdots, \psi_{1,2l}, \cdots, \psi_{2^k-1,0}, \cdots, \psi_{2^k-1,2l}]. \tag{14}$$

5 Operational Matrix of Fractional Calculus for Sine-Cosine Wavelet

In this section we find the operational matrix of fractional derivative for the considered wavelet using the operational matrix of fractional integration for BPF.

5.1 Express $\Psi(t)$ in Terms of BPF

$\psi_{n,m}(t)$ as a function can be express in terms of blockpulse function

$$\psi_{n,m} \simeq \sum_{i=0}^{m'-1} f_i b_i \quad m' = 2^k(2l+1), \tag{15}$$

$$f_i = m' \int_{\frac{i}{m'}}^{\frac{i+1}{m'}} \psi_{n,m}(x)dx = m' \int_{\frac{i}{m'}}^{\frac{i+1}{m'}} 2^{\frac{k+1}{2}} f_m(2^k x - n)dx. \tag{16}$$

Now we calculate f_i for different value of $i = 0, 1, \cdots, m'-1$

$$\begin{aligned} m = 0, \quad f_i = m' \int_{\frac{i}{m'}}^{\frac{i+1}{m'}} 2^{\frac{k+1}{2}} \times \frac{1}{\sqrt{2}} dx = 2^{\frac{k}{2}} \\ i = n(2l+1), \cdots, (n+1)(2l+1) - 1, \end{aligned} \tag{17}$$

$$\begin{aligned} m = 1, 2, \cdots, l, \quad f_i &= m' \int_{\frac{i}{m'}}^{\frac{i+1}{m'}} 2^{\frac{k+1}{2}} \cos(2m\pi(2^k x - n))dx \\ &= \frac{m'}{2^{\frac{k+1}{2}} m\pi} \left[\psi_{n,m+l}\left(\frac{i+1}{m'}\right) - \psi_{n,m+l}\left(\frac{i}{m'}\right) \right], \end{aligned} \tag{18}$$

$$\begin{aligned} m = l+1, \cdots, 2l, \quad f_i &= m' \int_{\frac{i}{m'}}^{\frac{i+1}{m'}} 2^{\frac{k+1}{2}} \sin(2(m-l)\pi(2^k x - n))dx \\ &= \frac{-m'}{2^{\frac{k+1}{2}}(m-l)\pi} \left[\psi_{n,m-l}\left(\frac{i+1}{m'}\right) - \psi_{n,m-l}(\frac{i}{m'}) \right]. \end{aligned} \tag{19}$$

For $m = 0$ we have

$$\psi_{n,m} = [\underbrace{0, \cdots, 0}_{n(2l+1)}, \underbrace{2^{k/2}, 2^{k/2}, \cdots, 2^{k/2}}_{2l+1}, 0, \cdots, 0] \times B_{m'}. \tag{20}$$

For $m = 1, 2, \cdots, l$

$$\begin{aligned} \psi_{n,m} = &\frac{m'}{2^{\frac{k+1}{2}} m\pi} \Bigg[\underbrace{0, 0, \cdots, 0}_{n(2l+1)}, \psi_{n,m+l}\left(\frac{n(2l+1)+1}{m'}\right) - \psi_{n,m+l}\left(\frac{n(2l+1)}{m'}\right), \\ &\cdots, \psi_{n,m+l}\left(\frac{(n+1)(2l+1)}{m'}\right) - \psi_{n,m+l}\left(\frac{n(2l+1)+2l}{m'}\right), 0, 0, \cdots, 0 \Bigg] \times B_{m'}. \end{aligned} \tag{21}$$

And for $m = l+1, \cdots, 2l$ we get

$$\begin{aligned} \psi_{n,m} = &\frac{-m'}{2^{\frac{k+1}{2}}(m-l)\pi} \Bigg[\underbrace{0, 0, \cdots, 0}_{n(2l+1)}, \psi_{n,m-l}\left(\frac{n(2l+1)+1}{m'}\right) - \psi_{n,m-l}\left(\frac{n(2l+1)}{m'}\right), \\ &\cdots, \psi_{n,m-l}\left(\frac{(n+1)(2l+1)}{m'}\right) - \psi_{n,m-l}\left(\frac{n(2l+1)+2l}{m'}\right), 0, 0, \cdots, 0 \Bigg] \times B_{m'}. \end{aligned} \tag{22}$$

Therefore we have $\Psi(x) = \Phi_{m' \times m'} B_{m'}(x)$ where $\Phi_{m' \times m'} = \text{diag}(\Phi_0, \Phi_1, \cdots, \Phi_{2^k-1})$, Φ_n is defined as follows, in the following matrix, $i = n(2l+1)$

$$\Phi_n = \begin{bmatrix} 2^{\frac{k}{2}} & \cdots\ 2^{\frac{k}{2}} \\ \frac{m'}{2^{\frac{k+1}{2}} 1\pi}(\psi_{n,1+l}\left(\frac{i+1}{m'}\right) - \psi_{n,1+l}(\frac{i}{m'}) & \cdots\ \psi_{n,1+l}(\frac{i+2l+1}{m'}) - \psi_{n,1+l}(\frac{i+2l}{m'})) \\ \vdots & \ddots\ \vdots \\ \frac{m'}{2^{\frac{k+1}{2}} l\pi}(\psi_{n,2l}\left(\frac{i+1}{m'}\right) - \psi_{n,21}\left(\frac{i}{m'}\right) & \cdots\ \psi_{n,2l}(\frac{i+2l+1}{m'}) - \psi_{n,21}(\frac{i+2l}{m'}))) \\ \frac{-m'}{2^{\frac{k+1}{2}} 1\pi}(\psi_{n,1}\left(\frac{i+1}{m'}\right) - \psi_{n,1}(\frac{i}{m'}) & \cdots\ \psi_{n,1}(\frac{i+2l+1}{m'}) - \psi_{n,1}(\frac{i+2l}{m'})) \\ \vdots & \ddots\ \vdots \\ \frac{-m'}{2^{\frac{k+1}{2}} l\pi}(\psi_{n,l}\left(\frac{i+1}{m'}\right) - \psi_{n,l}(\frac{i}{m'}) & \cdots\ \psi_{n,l}(\frac{i+2l+1}{m'}) - \psi_{n,l}(\frac{i+2l}{m'})) \end{bmatrix}. \tag{23}$$

5.2 Operational Matrix of Fractional Integration and Derivative for Sine-Cosine Wavelet

For finding operational matrix of fractional derivative of vector $\Psi(t)$, first of all we try to find the operational matrix of fractional integration.

$$(I^\alpha \Psi)(x) \simeq P^\alpha \Psi(x), \tag{24}$$

where P^α is the operational matrix of fractional integration, which calculate as follows

$$\begin{aligned} I^\alpha \Psi(x) &= I^\alpha \Phi_{m' \times m'} . B_{m'}(x) = \Phi_{m' \times m'} I^\alpha B_{m'}(t) = \Phi_{m' \times m'} F^\alpha B_{m'}(x) && (25) \\ &\Rightarrow P^\alpha \Psi(x) = P^\alpha \Phi_{m' \times m'} B_{m'}(x) = \Phi_{m' \times m'} F^\alpha B_{m'}(x) \\ &\Rightarrow P^\alpha = \Phi_{m' \times m'} F^\alpha \Phi^{-1}_{m' \times m'} \\ &\Rightarrow I^\alpha \Psi(x) \simeq \Phi_{m' \times m'} F^\alpha \Phi^{-1}_{m' \times m'} \Psi(x). && (26) \end{aligned}$$

Now we calculate operational matrix of derivative using P^α

$$D^\alpha f(x) = I^{n-\alpha} f^n(x) n-1 < \alpha \leq n n \in N, \tag{27}$$

$$D^\alpha x(t) = D^\alpha X^T \Psi(t) = X^T D^\alpha \Psi(t) = X^T I^{n-\alpha} \Psi^{(n)}(t). \tag{28}$$

For $\alpha \in$(0,1) we have $n = 1$ thus

$$\begin{aligned} D^\alpha x(t) &\simeq X^T I^{1-\alpha} D \Psi(t) = X^T D I^{1-\alpha} \Psi(t) \\ &= X^T D \Phi_{m' \times m'} F^{1-\alpha} \Phi^{-1}_{m' \times m'} \Psi(t), \end{aligned} \tag{29}$$

where D is operational matrix of derivative for $\Psi(t)$ which defined as $D = \text{diag}(w, w, \cdots, w)$, which is $2^k(2l+1) \times 2^k(2l+1)$ matrix and w is of size $(2l+1) \times (2l+1)$

$$w = 2^{k+1}\pi \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & -2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & -l \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & l & 0 & 0 & \cdots & 0 \end{bmatrix}_{(2l+1)\times(2l+1)} . \tag{30}$$

6 Solution of the Fractional Optimal Control Problem by Sine-Cosine Operational Matrix

Consider the fractional optimal control problem with quadratic performance index

$$\min J = \frac{1}{2}X^T(1)SX(1) + \frac{1}{2}\int_0^1 (X^T(t)QX(t) + U^T(t)RU(t))dt \tag{31}$$

$$\text{s. t.} \quad D^\alpha X(t) = AX(t) + BU(t) \tag{32}$$

$$X(0) = X_0 \quad 0 < \alpha \leq 1, \tag{33}$$

where A and B are constant matrices with the appropriate dimensions, also in cost functional S and Q are symmetric positive semi-definite matrices and R is a symmetric positive definite matrix. In this section, the Sine-Cosine wavelet is used for solving the above problem. We approximate each $x_i(t)$ and $u_i(t)$, in terms of Sine-Cosine wavelets as

$$X(t) = [x_1(t), x_2(t), \cdots, x_s(t)]^T \quad x_i(t) = \Psi^T(t)X_i \text{ or } X_i^T\Psi(t), \tag{34}$$

$$X(t) = \hat{\Psi}_s^T(t)X \quad X = [X_1^T, X_2^T, \cdots, X_s^T] \quad \hat{\Psi}_s(t) = I_s \otimes \Psi(t), \tag{35}$$

$$U(t) = [u_1(t), u_2(t), \cdots, u_q(t)]^T \quad u_i(t) = \Psi^T(t)U_i \text{ or } U_i^T\Psi(t), \tag{36}$$

$$U(t) = \hat{\Psi}_q^T(t)U \quad U = [U_1^T, U_2^T, \cdots, U_q^T] \quad \hat{\Psi}_s(t) = I_s \otimes \Psi(t), \tag{37}$$

where X_i, U_i are vectors of order $2^k(2l+1)\times 1$, X and U are vectors of order $s2^k(2l+1)\times 1$ and $q2^k(2l+1)\times 1$ respectively. $\otimes$ denotes the kronecker product. By substituting the above mentioned relation into objective function

$$J = \frac{1}{2}X^T\hat{\Psi}_s(1)S\hat{\Psi}_s^T(1)X + \frac{1}{2}\int_0^1 [X^T\hat{\Psi}_sQ\hat{\Psi}_s^TX + U^T\hat{\Psi}_qR\hat{\Psi}_q^TU]dt. \tag{38}$$

Since considered wavelet is orthonormal, it means $\int_0^1 \Psi^T(t)\Psi(t)dt = I$, we can rewrite Eq. (38) as follows

$$J(X,U) = \frac{1}{2}X^T[S\otimes\hat{\Psi}(1)\hat{\Psi}^T(1)]X + \frac{1}{2}[X^T(Q\otimes I)X + U^T(R\otimes I)U]. \tag{39}$$

Similarly, we do the same method for Eq. (32)

$$X(t) = X^T I_s \otimes \Psi(t) \text{ or } (I_s \otimes \Psi^T(t))X, \tag{40}$$

$$\begin{aligned} D^\alpha X(t) &= I^{1-\alpha} X'(t) = I^{1-\alpha}(X^T(I_s \otimes \Psi(t)))' = X^T I^{1-\alpha}(I_s \otimes (D\Psi(t)))' \\ &= X^T I_s \otimes (I^{1-\alpha} D\Psi(t)) = X^T I_s \otimes [D I^{1-\alpha}(\Psi(t))] \\ &= X^T I_s \otimes (D\Phi_{m'\times m'} F^\alpha \Phi^{-1}_{m'\times m'} \Psi(x)) \end{aligned} \tag{41}$$

$$\begin{aligned} R(t) &= X^T I_s \otimes (D\Phi_{m'\times m'} F^\alpha \Phi^{-1}_{m'\times m'})\Psi(t) - AX^T I_s \otimes \Psi(t) - BU^T I_q \otimes \Psi(t) R(t) \\ &= [X^T I_s \otimes (D\Phi_{m'\times m'} F^\alpha \Phi^{-1}_{m'\times m'}) - AX^T I_s \otimes I_{2^k(2l+1)} \\ &\quad - BU^T I_q \otimes I_{2^k(2l+1)}] \otimes \Psi(t). \end{aligned} \tag{42}$$

As in a typical tau method [5] we generate $2^k(2l+1)-1$ linear equations by applying

$$\langle R(t), \psi_{n,m}(t)\rangle = \int_0^1 R(t).\psi_{n,m}(t)dt = 0. \tag{43}$$

Also, by substituting Eq. (35) in (33) we get

$$X(0) = X^T \hat{\Psi}(0) = X_0. \tag{44}$$

Equations (43) and (44) generate $2^k(2l+1)$ set of linear equations. These linear equations can be solved for unknown coefficients of the vectors X^T and U^T. Consequently, $X(t)$ and $U(t)$ can be calculated.

7 Illustrative Example

We applied the method presented in this paper and solved the undergoing example.

Example 1. Consider the following time invariant FOCP [14],

$$\begin{aligned} \min J =& \frac{1}{2}\int_0^1 (x^2(t) + u^2(t))dt \\ \text{s. t.}\ \ & D^\alpha x(t) = -x(t) + u(t) \\ & x(0) = 1. \end{aligned}$$

We want to find a control variable $u(t)$ which minimizes the quadratic performance index J. This problem is solved by proposed method with $\alpha = 1, m = 5$ and $n = 7$, the numerical value obtained for J is 0.1979, which is close to the exact solutions in the case $\alpha = 1(0.1929)$.

8 Conclusion

In this paper, we derive a numerical method for fractional optimal control based on the operational matrix for the fractional integration and differentiation. The procedure of constructing these matrices is summarized. An example is given to show the efficiency of method. The obtained matrices can also be used to solve problems such as fractional optimal control with delay. Moreover we could find these matrices using another set of orthogonal functions instead of BPFs, it seems if we use a set of continuous orthogonal function the numerical result will improve.

References

1. Agrawal OP (2008) A formulation and numerical scheme for fractional optimal control problems. J Vibr Control 14(9–10):1291–1299
2. Bagley RL, Torvik P (1983) A theoretical basis for the application of fractional calculus to viscoelasticity. J Rheol 27(3):201–210
3. Baillie RT (1996) Long memory processes and fractional integration in econometrics. J Econometrics 73(1):5–59
4. Bohannan GW (2008) Analog fractional order controller in temperature and motor control applications. J Vibr Control 14(9–10):1487–1498
5. Canuto C, Hussaini MY, Quarteroni A, Zang TA Jr (1988) Spectral methods in fluid dynamics. Annu Rev Fluid Mech 57(196):339–367
6. Dehghan M, Manafian J, Saadatmandi A (2010) The solution of the linear fractional partial differential equations using the homotopy analysis method. Z Naturforsch A 65(Z. Naturforsch):935–949
7. Dehghan M, Manafian J, Saadatmandi A (2010) Solving nonlinear fractional partial differential equations using the homotopy analysis method. Numer Methods Partial Differ Equ 26(2):448–479
8. He J (1998) Nonlinear oscillation with fractional derivative and its applications. In: International conference on vibrating engineering, Dalian, China, vol 98, pp 288–291
9. He J (1999) Some applications of nonlinear fractional differential equations and their approximations. Bull Sci Technol 15(2):86–90
10. Jafari H, Tajadodi H (2014) Fractional order optimal control problems via the operational matrices of bernstein polynomials. Upb Sci Bull 76(3):115–128
11. Kajani MT, Ghasemi M, Babolian E (2006) Numerical solution of linear integro-differential equation by using sine-cosine wavelets. Appl Math Comput 180(2):569–574
12. Li Y, Sun N (2011) Numerical solution of fractional differential equations using the generalized block pulse operational matrix. Comput Math Appl 62(3):1046–1054
13. Liu F, Agrawal OP et al (1999) Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications. Academic Press
14. Lotfi A, Dehghan M, Yousefi SA (2011) A numerical technique for solving fractional optimal control problems. Comput Math Appl 62(3):1055–1067
15. Momani S, Odibat Z (2007) Numerical comparison of methods for solving linear differential equations of fractional order. Chaos Solitons Fractals 31(5):1248–1255

16. Odibat ZM, Shawagfeh NT (2007) Generalized taylor's formula. Appl Math Comput 186(1):286–293
17. Panda R, Dash M (2006) Fractional generalized splines and signal processing. Sig Process 86(9):2340–2350
18. Rossikhin YA, Shitikova MV (1997) Applications of fractional calculus to dynamic problems of linear and nonlinear hereditary mechanics of solids. Appl Mech Rev 50:15–67
19. Saadatmandi A, Dehghan M (2010) A new operational matrix for solving fractional-order differential equations. Comput Math Appl 59(3):1326–1336
20. Saez D (2009) Analytical solution of a fractional diffusion equation by variational iteration method. Comput Math Appl 57(3):483–487
21. Sen S (2011) Fractional optimal control problems: A pseudo-state space approach. J Vibr Control 17(17):1034–1041
22. Shawagfeh NT (2002) Analytical approximate solutions for nonlinear fractional differential equations. Appl Math Comput 131(2–3):517–529
23. Sohrabi S (2011) Comparison chebyshev wavelets method with bpfs method for solving abel's integral equation. Ain Shams Eng J 2(3–4):249–254
24. Sweilam NH, Alajami TM, Hoppe RHW (2013) Numerical solution of some types of fractional optimal control problems. Sci World J 2013(2):306237
25. Tangpong XW, Agrawal OP (2009) Fractional optimal control of continuum systems. J Vibr Acoust 131(2):557–557

The Sustainable Interaction Analysis of Cause Marketing and Ethical Consumption in Electric Business Platform: Based on Game Theory and Simulation Analysis

Xu Zu and Weiping Yu(✉)

Business School, Sichuan University, Chengdu 610065, People's Republic of China
331382107@qq.com

Abstract. Based on evolutionary game theory and simulation analysis, the paper studies the sustainable interaction of cause marketing in electronic business platform, which constructs the evolutionary game model of cause marketing and ethical consumption, and studies the existence and stability of the equilibrium in the four situation interaction system, and also verifies results by Netlogo. The main contributions are that ethical consumption is the precondition of sustainable cause marketing development, reducing costs is the key to promote sustainable interaction development, electric business platform should vigorously promote social responsibility consciousness, and promote the policy support. Finally the paper constructs the theory frame for the realization of benign sustainable interaction.

Keywords: Cause marketing · Ethical consumption · The evolutionary game model · Simulated analysis

1 Introduction

In recent years, the development rate of China's e-commerce is 2–3 times of GDP (7% 9%). China's B2C online shopping market deals reached 609.67 billion yuan in the second quarter of 2016. With the rise of e-commerce, the businessman tried to introduce a cause marketing methods to get more customers focus on their homogeneous products, for example, some businesses join the public welfare plan of Taobao platform. Cause marketing dated from the public welfare activity of "Renovation of the statue of liberty" by the cooperation of American Express Company and Alice Island Foundation, and then it obtained the widespread attention and development. Nowadays, more and more traditional corporations have been successfully performed the corporate social responsibility by the form of cause marketing. The cause marketing is not only contributing to society, but also obtaining the business interests [1,8], for example, the cause marketing activity of "buy a bottle of water, donate a penny" improved Nongfu

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_19

spring's product sales, and Nongfu Spring raised more than 5 million yuan, 361° "ONE CARES ONE" also promoted sales of especially-made shoes and expanded brand influence. Cause marketing, also known as charity Marketing or cause-related Marketing, is a special form of donation. It depends on the purchasing behavior of consumers, and donates the amount of a certain percentage of business to the special public welfare project. Cause marketing has become one of the important way of charity enterprises to perform social responsibility in many countries. Cause marketing can promote sales, social responsibility image and brand awareness [9,13]. In E-commerce platform, corporations, which successfully continue to push the cause marketing, not only can achieve the aim of marketing businesses, but also can effectively promote the healthy development of Chinese social and economic. However, focusing on E-commerce situation, businesses joined the cause marketing activities in E-commerce platform, and didn't get a good response of consumers.

From the practice experience of western countries, the benign interaction between the enterprise and the consumer is the key to successfully propel the corporate social responsibility [12]. The concept of ethical consumption is that consumers not only consider the commercial value of product, but also consider the beneficial impact of their purchasing behavior on society, environment, and so on [5], that effectively promote the corporate cause marketing to fulfill the social responsibility better. The sustained interaction of cause marketing and ethical consumption can ensure that businessmen get more consumer approval ratings in order to cover cost and promote sales and corporate image. According to the survey in U.S. market, 78% of consumers are willing to buy products with cause marketing, 66% of consumers willing to switch brands to support cause marketing [3], and 96.6% consumers also incline to choose homogeneous products with good corporate social image [6] in China market. However, consumers are also "rational economic man". Because of the higher perceived risk in the E-commerce platform, consumers are not willing to fulfill the ethical consumption with much cost in order to balance the benefit and the cost [11], so there is often the inconsistent phenomenon about the consumer ethical attitude and behavior.

The relationship researches of cause marketing and ethical consumption mainly focus on two aspects: the first is the relationship of cause marketing and ethical consumption based on the investigation method and experiment method [1,9], the second is the cause of the inconsistent about consumers' ethical attitude and behavior based on the interview and other qualitative methods [2]. However, few scholars' researches focus on the interaction mechanism, especially the research of systemic balance and dynamic evolution in the E-commerce platform. It belongs to a kind of gambling behavior that cause marketing and ethical consumption. This research constructs the evolutionary game model based on evolutionary game theory and systematically explores the evolution rules. The research result not only enriches the existing theory research of cause marketing and ethical consumption, but also provides valuable advice on sustainable promoting the cause marketing and ethical consumption in the E-commerce platform.

2 Evolutionary Game Model Building

Evolutionary game theory emphasizes the dynamic equilibrium. It can real reaction behavior diversity and complexity of the body combined the game theory analysis and the dynamic evolution process analysis [7,10]. Based on evolutionary game theory, this research adopts the method that two kinds of group repeated game and replication dynamic evolutionary game, and puts forward the evolutionary game model to research like the assumptions:

Hypothesis 1: The two types of players are businesses and consumers in the E-commerce platform, and they are the rational "economic individuals". The businesses behavior sets are cause marketing and non cause marketing and the consumers behavior sets are ethical consumption and unethical consumption.

Hypothesis 2: The proportion of businesses with cause marketing is x, and the other is $1 - x$; the proportion of consumers with ethical consumption is y, and the other is $1 - y$.

Hypothesis 3: The product is homogeneous, the product cost for cause marketing is C_1 and product sales price is P_1, The endogenous utility, the brand image and perceived social contribution and so on, and the $E-$ commerce platform support utility are F; the product cost for non cause marketing is C_2 and product sales price is P_2, and C_1 is higher than C_2.

Hypothesis 4: The social responsibility consciousness of consumers is strengthened with the development of society, the ethical consumption not only meets the basic demand for products, but also makes consumers additional utility of social contribution, the utility is V, and K is the additional cost that consumers need to pay higher for ethical consumption than unethical consumption.

Based on the above assumptions, businesses' revenue is as follows:

(1) When businesses adopt cause marketing and consumers adopt ethical consumption, businesses' revenue is $P_1 - C_1 + F$;
(2) When businesses adopt cause marketing and consumers adopt unethical consumption, businesses' revenue is $-C_1 + F$;
(3) When businesses adopt non cause marketing and consumers adopt ethical consumption, businesses' revenue is $-C_2$;
(4) When businesses adopt non cause marketing and consumers adopt unethical consumption, businesses' revenue is $P_2 - C_2$.

Consumers' revenue is as follows:

(1) When consumers adopt ethical consumption and businesses adopt cause marketing, consumers' revenue is $V - K$;
(2) When consumers adopt ethical consumption and businesses adopt non cause marketing, consumers' revenue is $-V$;
(3) When consumers adopt unethical consumption and businesses adopt cause marketing or non cause marketing, consumers' revenue is 0. Payoff matrix for gaming revenue are shown in Table 1.

Table 1. Payoff matrix for gaming revenue about businesses and consumers.

	Type	Consumers	
		Ethical consumption (y)	Unethical consumption $(1-y)$
Businesses	Cause marketing (x)	$P_1 - C_1 + F, V - K$	$-C_1 + F, 0$
	Non cause marketing $(1-x)$	$-C_2, -V$	$P_2 - C_2, 0$

Expected returns of cause marketing and non cause marketing respectively are U_0^1, U_0^2, the average return of businesses is $\overline{U}_0$:

$$U_0^1 = y(P_1 - C_1 + F) + (1-y)(-C_1 + F) = yP_1 - C_1 + F, \tag{1}$$

$$U_0^2 = y(-C_2) + (1-y)(P_2 - C_2) = (1-y)P_2 - C_2, \tag{2}$$

$$\overline{U}_0 = xU_0^1 + (1-x)U_0^2. \tag{3}$$

Expected returns of ethical consumption and unethical consumption respectively are U_1^1, U_1^2, the average return of consumers is $\overline{U}_1$:

$$U_1^1 = x(V-K) + (1-x)(-V) = x(2V-K) - V, \tag{4}$$

$$U_1^2 = x \cdot 0 + (1-x) \cdot 0 = 0, \tag{5}$$

$$\overline{U}_1 = yU_1^1 + (1-y)U_1^2. \tag{6}$$

According to the above expressions, we can know the replicated dynamic equations respectively are:

$$\dot{x} = \frac{dx}{dt} = x(U_0^1 - \overline{U}_0) = x(1-x)\left[y(P_1 + P_2) + F - C_1 - P_2 + C_2\right], \tag{7}$$

$$\dot{y} = \frac{dy}{dt} = y(U_1^1 - \overline{U}_1) = y(1-y)\left[x(2V-K) - V\right]. \tag{8}$$

3 Evolutionary Game Model Analysis

3.1 Equilibrium Point and Stability Analysis

Make $\frac{dx}{dt} = 0$ and $\frac{dy}{dt} = 0$, based on results of replicated dynamic equation of the interaction between the businesses and the consumers, we can get five local equilibrium points: $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$ and (p^*, q^*), and

$$p^* = \frac{V}{2V-K}, \tag{9}$$

$$q^* = \frac{C_1 - F + P_2 - C_2}{P_1 + P_2}. \tag{10}$$

Friedman proposed that the stability of the equilibrium point of the evolution system can be obtained from the local stability analysis of the jacobian matrix (denoted by J) [4]. The jacobian matrix by the combination of Eqs. (7) and (8):

$$J = \begin{pmatrix} \frac{\partial \dot{x}}{\partial x} & \frac{\partial \dot{x}}{\partial y} \\ \frac{\partial \dot{y}}{\partial x} & \frac{\partial \dot{y}}{\partial y} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \tag{11}$$

and

$$a_{11} = (1 - 2x)\left[y(P_1 + P_2) + F - C_1 - P_2 + C_2\right], \tag{12}$$

$$a_{12} = x(1 - x)(P_1 + P_2), \tag{13}$$

$$a_{21} = y(1 - y)(2V - K), \tag{14}$$

$$a_{22} = (1 - 2y)\left[x(2V - K) - V\right]. \tag{15}$$

Thus, we can get the numerical results about a_{11}, a_{12}, a_{21} and a_{22} under the five local equilibrium points, the results are shown in Table 2.

Table 2. The numerical results about a_{11}, a_{12}, a_{21} and a_{22} under the five local equilibrium points.

Equilibrium point	a_{11}	a_{12}	a_{21}	a_{22}
$(0, 0)$	$F - C_1 - P_2 + C_2$	0	0	$-V$
$(1, 0)$	$-F + C_1 + P_2 - C_2$	0	0	$V - K$
$(0, 1)$	$P_1 + F - C_1 + C_2$	0	0	V
$(1, 1)$	$-P_1 - F + C_1 - C_2$	0	0	$K - V$
(p^*, q^*)	0	A	B	0

And the expressions of A and B respectively are:

$$A = \frac{V}{2V - K}(1 - \frac{V}{2V - K})(P_1 + P_2), \tag{16}$$

$$B = \frac{C_1 - F + P_2 - C_2}{P_1 + P_2}(1 - \frac{C_1 - F + P_2 - C_2}{P_1 + P_2})(2V - K). \tag{17}$$

When the jacobian matrix satisfies the following conditions of the determinant value (denoted by $\det J$) and trace value (denoted by trJ), the local equilibrium of system is the stability and can be the evolutionary stable strategy (denoted by ESS).

Condition 1:

$$\det J = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} > 0, \tag{18}$$

Condition 2:

$$trJ = a_{11} + a_{22} < 0. \tag{19}$$

Obviously, when the local equilibrium of system is (p^*, q^*), $a_{11} + a_{22} = 0$ do not meet condition 2, so it is not as ESS. The study needs to discuss the possibility that four local equilibrium points become ESS from different situations in the plane of $N = \{x, y\,|0 \le x, y \le 1\}$.

3.2 Result Discussion

Case 1: When $F > C_1 + P_2 - C_2$ and $V > K$, the stability of four local equilibrium points is as the Table 3 shown. We can get the ESS when $(1, 1)$ satisfies the conditions of $\det J > 0$ and $trJ < 0$.

Table 3. The stability analysis of local equilibrium point when $F > C_1 + P_2 - C_2$ and $V > K$.

Equilibrium point	$detJ$	trJ	Stability
$(0, 0)$	$-$	Indeterminacy	Saddle point
$(1, 0)$	$-$	Indeterminacy	Saddle point
$(0, 1)$	$+$	$+$	Instability point
$(1, 1)$	$+$	$-$	ESS

Case 2: When $F < C_1 + P_2 - C_2$ and $V > K$, the stability of four local equilibrium points is as the Table 4 shown, We can get the ESS when (0,0) and (1,1) satisfy the conditions of $\det J > 0$ and $trJ < 0$.

Table 4. The stability analysis of local equilibrium point when $F < C_1 + P_2 - C_2$ and $V > K$.

Equilibrium point	$detJ$	trJ	Stability
$(0, 0)$	$+$	$-$	ESS
$(1, 0)$	$+$	$+$	Instability point
$(0, 1)$	$+$	$+$	Instability point
$(1, 1)$	$+$	$-$	ESS

Case 3: When $F > C_1 + P_2 - C_2$ and $V < K$, the stability of four local equilibrium points is as the Table 5 shown, We can get the ESS when (1, 0) satisfy the conditions of $\det J > 0$ and $trJ < 0$.

Case 4: When $F < C_1 + P_2 - C_2$ and $V < K$, the stability of four local equilibrium points is as the Table 6 shown, We can get the ESS when (1,0) satisfy the conditions of $\det J > 0$ and $trJ < 0$.

Table 5. The stability analysis of local equilibrium point when $F > C_1 + P_2 - C_2$ and $V < K$.

Equilibrium point	$detJ$	trJ	Stability
$(0, 0)$	$-$	Indeterminacy	Saddle point
$(1, 0)$	$+$	$-$	ESS
$(0, 1)$	$+$	$+$	Instability point
$(1, 1)$	$-$	Indeterminacy	Saddle point

Table 6. The stability analysis of local equilibrium point when $F < C_1 + P_2 - C_2$ and $V < K$.

Equilibrium point	$detJ$	trJ	Stability
$(0, 0)$	$+$	$-$	ESS
$(1, 0)$	$-$	Indeterminacy	Saddle point
$(0, 1)$	$+$	$+$	Instability point
$(1, 1)$	$-$	Indeterminacy	Saddle point

4 Netlogo Simulation Modeling and Analysis

The study validates and analyzes the stability of equilibrium points with the aid of Netlogo modeling and simulation in the four situations. Netlogo is developed by the center of connectionist learning and computer modeling in northwestern university. It is based on multi-agent modeling, multi-agent parallel and asynchronous updating, and the whole system is dynamic with time. It can well study the macroscopic patterns of micro-individual interaction, and is especially suitable for modeling and simulating complex systems with time evolution. Based on the analysis of the stability of the equilibrium point in the evolutionary game model, this study sets the payment matrix from four cases and performs simulation verification and analysis. The horizontal axis represents the simulation time and the vertical axis represents the ratio that the number of subjects to the total number of game strategy in the simulation results chart. $X_1 Rate$ and $Y_1 Rate$ respectively are the cause marketing rate and ethical consumption rate, and the initial ratio is 1/2.

Case 1: From the stability analysis result of the replicated dynamic equation, (1,1) is ESS when $F > C_1 + P_2 - C_2$ and $V > K$. The study substitutes $F = 8$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$ and $K = 1$ into the simulation model, and the simulation result is as shown in Fig. 1.

The simulation results validate the deduction of the replicated dynamic equation. The evolution result is an ideal phenomenon that businesses, consumers and electric business platform are involved in social responsibility behavior. Consumers need to pay additional costs, but the additional utility of social contribution is always greater than the additional costs paid, consumers still choose ethical consumption strategy; cause marketing strategy will increase the cost of income, but a strong sense of social responsibility and good support and encouragement of electric business platform will increase the business additional benefits, so businesses still prefer to cause marketing strategy. Responsibility behavior of businesses and consumers will enter a good time of continuous interactive development.

Case 2: From the stability analysis result of the replicated dynamic equation, (1,1) and (0,0) are ESS when $F < C_1 + P_2 - C_2$ and $V > K$. The study respectively substitutes $F = 4$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$ and $K = 1$ and $F = 1$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$ and $K = 0.5$ into the simulation model, and the simulation result is as shown in Figs. 2(a) and 3(a).

The simulation results validate that there are two kinds of behavior patterns in the evolution system, namely [cause marketing, ethical consumption] and [non cause marketing, unethical consumption]. When the benefits from a good corporate brand image and self-perception of social contributions and commercial policy benefits make up for the loss of profits due to cause marketing, businesses will still choose cause marketing. Consumers, compensated for the additional cost, will still choose the ethical consumption strategy, so businesses and consumers are in a benign behavioral interaction process. When the additional benefits of cause marketing can not make up for the loss of profits, the system will quickly move toward a vicious direction: businesses will choose non cause marketing strategy, consumers will choose unethical consumption strategy. Social responsibility awareness of businesses and consumers is low and the system fails to achieve the sustainable social interaction due to lack of awareness and support of electric business platform.

Case 3: From the stability analysis result of the replicated dynamic equation, (1,0) is ESS when $F > C_1 + P_2 - C_2$ and $V < K$. The study substitutes $F = 8$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$ and $K = 4$ into the simulation model, and the simulation result is as shown in Fig. 4.

The simulation results validate that businesses will still keen on cause marketing although they can not sell products when the additional effect is greater than the sum of non cause marketing profits and the cause marketing cost. Consumers, as "economic man", can not pay too much extra expense for ethical consumption, and finally give up ethical consumption. The stability of system is based on policy support of the electric business platform and businesses own economic interests, if the electronic business platform reduce support and the economic benefits can not make up for cause marketing cost, then the system will be rapidly unbalanced.

Case 4: From the stability analysis result of the replicated dynamic equation, (0,0) is ESS when $F < C_1 + P_2 - C_2$ and $V < K$. The study substitutes $F = 1$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$ and $K = 4$ into the simulation model, and the simulation result is as shown in Fig. 5.

The simulation results validate that consumers can not use the extra utility of ethical consumption to make up for the extra cost, so they abandon the ethical consumption, and gradually choose unethical consumption through game learning. Although businesses get a certain income because of cause marketing,

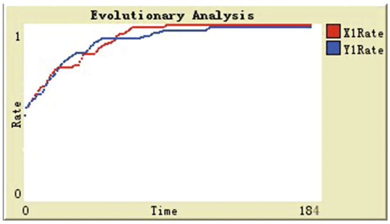


Fig. 1. The simulation result ($F = 8$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$, $K = 1$)

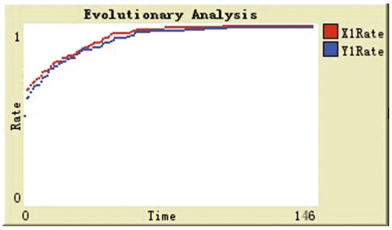


Fig. 2. The simulation result ($F = 4$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$, $K = 1$)

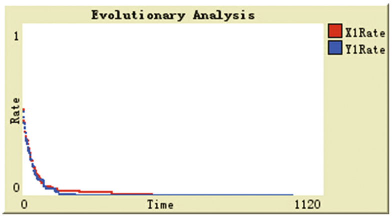


Fig. 3. The simulation result ($F = 1$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$, $K = 0.5$)

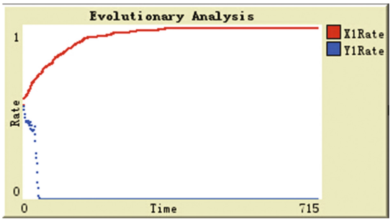


Fig. 4. The simulation result ($F = 8$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$, $K = 4$)

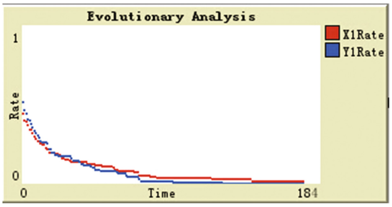


Fig. 5. The simulation result ($F = 1$, $C_1 = 4$, $P_2 = 5$, $C_2 = 2$, $V = 3$, $K = 4$)

businesses finally take non cause marketing strategy in order to sell products for profit without ethical consumption support, so the system will evolve to a vicious direction, and will be rapidly unbalanced.

5 Conclusion and Suggestion

From the perspective of evolutionary game, this paper uses replication dynamic evolutionary game method and Netlogo numerical simulation to analyze the sustainable interaction mechanism between cause marketing and ethical consumption, and obtains the positive and sustainable interaction conditions of cause marketing and ethical consumption.

(1) Ethical consumption is a prerequisite for the sustainable development of cause marketing. It's the safeguard for the sustainable development of cause marketing that consumers are willing to ethical consumption. Consumer cost affects choices of consumer behavior, low level of social responsibility awareness will directly increase the support burden of electric business platform to cause marketing, can not guarantee the product sales and achieve the good interaction of cause marketing and ethical consumption. Therefore, electronic business platform and businesses should implement the cause marketing, increase the cause marketing publicity to enhance consumer awareness of social responsibility, while reducing the time cost that consumer understand the specific content of cause marketing.

(2) Reducing cost is the key to promote sustainable development of cause marketing. Cause marketing cost is directly related to the sustainable interactive development of cause marketing and ethical consumption. Businesses should work from two aspects: First, reduce the cause marketing cost. The price of cause marketing product is not necessarily the same as the price of non cause marketing product, but it can not exceed the maximum afforded range of ethical consumption. The price should be combined with consumer's sense of social responsibility, at the same time, businesses should use other ways to reduce cause marketing cost. Second, businesses should indirectly reduce consumer ethical consumption cost. Businesses should accurately analyze consumer behavior, so as to accurately put cause marketing for advertising, really more convenient for consumers to understand the specific content of cause marketing, thereby reducing consumer spending on the time and other additional costs, better to achieve sustainable interaction of cause marketing and ethical consumption.

(3) Electronic business platform should vigorously promote the sense of social responsibility and promote the platform policy support. Businesses, as "economic man", always want their own additional endogenous utility and electric platform support for cause marketing to make up for opportunity cost and accounting cost. The platform should guide businesses to enhance the sense of social responsibility, so that the maximum expectation of policy support defined in the reasonable range of C1–C2. The specific work can be implement from two aspects: First, increase social responsibility propaganda. Only relying on the platform support policy is unable to maintain the sustainable development of cause marketing. Its important to increase the effectiveness of business and consumer social responsibility that improve the social responsibility sense in the electric business platform. Electric business platform should actively promote social responsibility concept to improve social responsibility awareness of businesses and consumers in order to increase the effectiveness of social responsibility. Second, improve the social responsibility incentive mechanism. Cause marketing will increase the cost of investment, electric business platform should speed up the development of policy incentives and other business incentives, to a certain extent make up the cost burden of business social responsibility, to better promote the sustainable development of cause marketing.

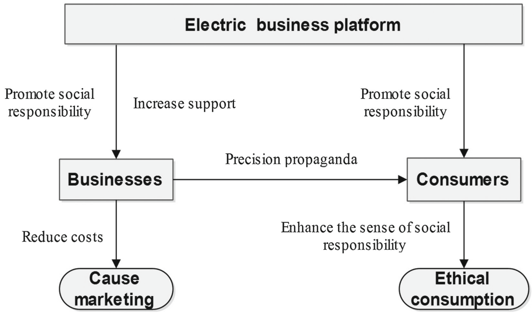


Fig. 6. The theoretical framework for the realization of benign and sustainable interaction.

Based on the above conclusions, this paper constructs a theoretical framework for the realization of benign and sustainable interaction between cause marketing and ethical consumption, as shown in Fig. 6:

Acknowledgements. This paper is supported by the General Project of the National Natural Science Foundation of China “The Mechanism of Effect of Corporate Social Responsibility Behavior on the Image of Scandal Brand and Strategy Research” (No. 71372189).

References

1. Al-Dmour H, Al-Madani S et al (2016) Factors affecting the effectiveness of cause-related marketing campaign: moderating effect of sponsor-cause congruence. Int J Mark Stud 8(5):114–127
2. Carrigan M, Attalla A (2001) The myth of the ethical consumer - do ethics matter in purchase behaviour? J Consum Mark 18(7):560–578
3. Chang CT (2008) To donate or not to donate? Product characteristics and framing effects of cause-related marketing on consumer purchase behavior. Psychol Mark 25(12):1089–1110
4. Friedman D (1991) Evolutionary games in economics. Econometrica 59(3):637–666
5. Hoffmann S, Hutter K (2012) Carrotmob as a new form of ethical consumption. The nature of the concept and avenues for future research. J Consum Policy 35(2):215–236
6. Huo B, Zhou Y (2014) Research on relationship of corporate social responsibility, corporate reputation and corporate performance. J Ind. Technol Economics 1:59–65 (in Chinese)
7. Khalifa NB, El-Azouzi R et al (2016) Evolutionary games in interacting communities. Dynamic Games Appl 1:1–26
8. Kim YJ, Lee WN (2009) Overcoming consumer skepticism in cause-related marketing: the effects of corporate social responsibility and donation size claim objectivity. J Promot Manage 15(4):465–483
9. Plewa C, Conduit J, Quester PG (2015) The impact of corporate volunteering on CSR image: a consumer perspective. J Bus Ethics 3:643–659

10. Selten R (1986) Evolutionary stability in extensive two-person games - correction and further development. Math Soc Sci 3:223–266
11. Shaw D, Mcmaster R, Newholm T (2016) Care and commitment in ethical consumption: an exploration of the ‘attitude-behaviour gap’. J Bus Ethics 136(2):251–265
12. Shin Y, Thai VV (2015) The impact of corporate social responsibility on customer satisfaction, relationship maintenance and loyalty in the shipping industry. Corp Soc Responsib Environ Manage 6:381–392
13. Vanhamme J, Lindgreen A et al (2012) To do well by doing good: improving corporate image through cause-related marketing. J Bus Ethics 109(3):259–274

Data Analysis

How to Predict Financing Efficiency in Public-Private Partnerships–In an Aspect of Uncertainties

Yixin Qiu[1], Umair Akram[1](✉), Sihan Lin[1], and Muhammad Nazam[2]

[1] School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
leo_umairinfo@yahoo.com

[2] Institute of Business and Management Science, University of Agriculture Faisalabad, Faisalabad, Pakistan

Abstract. Financing efficiency in Public-Private Partnership (PPP) projects is a vital but impressionable aspect, whose likely consequence always influenced by inevitable uncertainties and confused the stakeholders. This study aims to investigate the methods for evaluating and predicting the relative efficiency of current financing scheme in PPP project, which surrounding with various uncertainties. The current study proposed a new algorithm for assessing the inancing efficiency and simulating the flexibilities in PPP projects simultaneously. Furthermore, it integrates DEA and Monte Carlo to provide the possible probability distribution of financing efficiency which takes the uncertainties in PPP. A case study is given to show its feasibility and practicability. It shows uncertainties in PPP projects considerably count for its deviation from the optimum value. This innovative model provides an avenue to predict the effects and possible outcome of an optimal scheme, which is important for financing scheme decision making.

Keywords: PPP · Monte Carlo · DEA · Financing efficiency

1 Introduction

Infrastructure constructions are essential to support the world economic development. Both in developed and developing countries, governments have dedicated a significant share of the public budget to develop and refurbish infrastructures [10]. While, with the global economy slows, the lack of economic resource can prevent further economic growth and the development of already planned infrastructure projects [5]. Infrastructure builders (always public sectors) are driven to attract private investment to take part in infrastructure constructions. This idea has elicited the generation of Public-Private Partnerships (PPP) mode.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_20

The development and popularization of PPP well bridge the gap between a need of infrastructures and a lack of budget. Worldwide, PPP grows into one of the most important procurement mechanisms of project development.

PPP mode holds numerable advantages, among which the financial efficiency is one of the main superiority. The British government directly name PPP projects as “Private Finance Initiative”, therefore its effect on financing is evidently reflected. The Asian Development Bank strong introduced PPP for constructing infrastructures at the beginning of its establishment, in order to reach the value of money and promote the efficiency of financing [1]. Many scholars studied the financing effects of PPP mode either from theoretical aspect or relating with practical application. Sastoque [12] concluded that it is one of the key characteristics of PPP to count private sector’s contribution on financial issues of public projects. Gatti indicated PPP is an efficient financing instrument acts through designing, structuring, and executing, as an important role in whole life cycle [6]. The high financing efficiency of PPP mode allows infrastructure projects obtain the best value for money.

However, stakeholders of PPP are in need of a benchmark to learn about the financing efficiency quantitatively; it is a necessary prerequisite for making a sensible decision before it simplement. Uncertainties implicated in PPP projects, like lengthy durations, high transaction costs and a lack of competition and transparency, always lead to inefficiencies and ineffectiveness [2], and further drag the financing efficiency down. Experts are trying to build a systematic model for the assessment of financing efficiency in PPP. While, PPP projects are multi-participation activities and multi-indicator processes which makes this work more complicated than traditional ones, let alone making predictions and selecting the optimal scheme.

Various methods have been proposed by scholars to settle this issue. The first family of methods universally used in the assessment of PPP projects is Fuzzy Synthetic Evaluation (FSE). Verweij [14] employed FSE to analysis satisfactory outcome of financing in 27 PPP road constructions in the Netherlands. FSE performs well in comparing existing options, however, fails to propose an optimum scheme and make prediction or simulation. The second family of methods for addressing the above problem is Analytic Hierarchy Process (AHP) analysis. Zhuang [21] applied AHP to research the fiscal control in PPP of city infrastructure. AHP suffered from the same limitation as FSE. Besides, it is easily influenced by the subjective rating of experts. Recently, an improved method had been proposed for achieving a comparison of the existing options, optimization of the inefficient ones, as well as eliminating subjective influence from human. It is Data Envelopment Analysis (DEA). In YuYuanchun’s study, DEA plays an important role in evaluating the technology transfer efficiency in industry-university-research institution [17]. Wang Hong investigated the efficiency of debt risk and fiscal expenditure of local government with an application of DEA [15]. Another strong point of DEA is that it supports target data of each indicator in relatively inefficient projects, which provides a reference for optimizing the original scheme. But it ignores the uncertainties and external influences in projects implement, which can be conducted by Monte Carlo (MC) approach.

The objective of this work is to propose a mathematical model to evaluate and predict the financing efficiency in PPP Projects. We apply DEA method to handle difficulties caused by multiple participators and factors, and integrate MC method to simulate uncertainties embraced in PPP projects. It is may offer a new insight and academic reference for related research.

2 Problem Statement

The motivation of the proposed method lies in the fact that the crude FSE and AHP method are limited in giving general evaluation and selecting the optimal scheme; moreover, they are vulnerable to anthropic rating [9]. These restrictions result in a series of dilemma in PPP projects, like hard to assess the potential influence of uncertainties, and lack of quantitative support for decision making of financing scheme. So, stakeholders in PPP projects are in an urgent need of a systematic model that offers data support for the decision of financing strategies, which should be uncertainties considering, optimum proposal providing, and free from artificial factors.

DEA is able to achieve all the mentioned goals except that it fails to take uncertainties in the implement and operation of PPP projects into account. The principle of Monte Carlo analysis is to run stochastic values of independent variable sand then simulate and identifies the possible distribution of outcomes [20]. A massive repetition of simulation processes counts the flexibility of changing environment, which supplement the flaw of DEA.

In this work, we try to propound an approach by exploiting Monte Carlo approach and integrating it with DEA. Monte Carlo works for simulating the uncertainties in PPP projects which facilitate the new method to systematically fulfill evaluation and optimization of financing efficiency in PPP Projects.

3 Monte Carlo and Its Application

Monte Carlo (MC) simulation is a computerized mathematical technique that allows people to simulate an approximate solution in quantitative analysis and decision making. The technique was first exploited by scientists who work in the nuclear technology field when they were facing with a 6-dimensional integral equation [20]. Calculating a result by repeated simulation was the only way to solve this complex multiple problem. Since then, MC was widely applied in the nuclear area, however, be confined to this area also because of the restriction of the capability of computing. As in projects, the related influence factor can be even more complex compared with in nuclear technology, a strong ability in computing and storage is needed [4]. With the booming development of computer in recent years, MC is widespread to various fields, for handling the simulation of multidimensional problems, especially in project decision making which constantly face with uncertainty, ambiguity, and variability. By using this method, the distribution of all possible outcome of an event is generated by running a model a large scale of times, and then prompts and assists the decision-making [11]. The main principle of MC is simulating the input of multiple independent variables, in accordance with their expected distributions, and drawing the proximate distribution and statistic data of dependent index. In the procedures of MC, it always involves determining the change rules of independent factors and exploring their interaction with dependent outcome. Then, the running of numerous simulations could identify the range and distribution of possible outcomes according to a number of scenarios. While managing a project, administrators always face up with difficulties when there is a list of alternatives for the project but lack of clues of their possible impact on the project. MC technique helps in forecasting the likely outcome of an event and thereby promotes making an informed decision. The method algorithm is shown in its succession interactive five steps [4,11,20]:

Step 1. Clearly defining the target problem y, and creating a parametric model composed by its influence factors $x_1, x_2, \cdots, x_q$, the model is noted as Eq. (1):

$$y = f(x_1, x_2, \cdots, x_q). \tag{1}$$

Step 2. On the basis of fitted or hypothetic distributions of influence factors, generating a random input set of data, $x_1, x_2, \cdots, x_q$;

Step 3. Adding the random set which is generated from step 2 into the model in step 1, obtaining an effective calculation and recording results as y_i;

Step 4. According to the accuracy requirement of specific research, determining the simulation times n, and repeating steps 2 and 3 for $i = 1$ to $i = n$ ($n = 1000$ in this work);

Step 5. Obtaining a distribution curve and analyzing the results by using statistic indicators generated from the simulation.

In the execution of PPP projects, forecasting the final financing efficiency is a complex and repeated issue. It always bewilders stakeholders because of its high degree of complexity and uncertainty. MC simulation assists in simulating likely scenarios and giving a probable distribution of objectives with taking the influence of risk factors and uncertainties into consider. In the model, we put forward in this study, DEA is applied to take the place of parametric model included in step 1. The key benefits and applicability of using the MC analysis in this research are listed below:

- MC simulation settles the uncertainties caused by multiple participators and complex influence factors in PPP projects well. Differing from the crude traditional model, MC exempts the complex mathematical calculation, and directly obtains likely outcomes by simulating a series of possible scenarios, which make it more operable and practical.
- MC overcomes the irreversibility involved in decision making, especially in significant determination as those in PPP projects. Simulation is a way to imitate results by running a large number of potential situations. MC simulation meets the need of foresight the possible outcome in advance.

- With the advanced computing technology, Monte Carlo simulation is easier to operate. It gains the availability and feasibility in practical application of PPP projects.

4 Model Establishment

Scholars have explored a number of avenues to calculate and optimize the productivity and operational efficiency issues. However, it is noteworthy that DEA is one of the most popular approaches in this area. It is more recent in applications among studies and it also draws more effective conclusions in comparison with some other methodologies like Stochastic Frontier Analysis (SFA) [13]. With a popularization of the PPP mode, employing DEA as modeling method is becoming more and more popular in latest decade [19]. While in the specific financing efficiency field of PPP projects, only a few scholars have gone into and DEA is scarcely used.

In this research, we first adopt DEA to evaluate the financing efficiency of specific PPP projects and get the initially optimized data. Due to the limitation of theoretical mathematical models, DEA ignores the uncertainties and all possible scenarios in PPP projects. Therefore, Monte Carlo is designed to play a role for filling in the vacancy. The integrated algorithm can be used to both give an assessment and practically optimized scheme and its efficiency.

4.1 Evaluation and Optimization of Financing Efficiency in PPP Projects

In the projects area, the application of DEA always embodies three main parts: selecting input and output indexes, forming the decision-making unit set, and finally analyzing from calculated outcome. For the efficiency of financing in PPP projects, it involves two major areas: raising funds in low cost and using funds efficiently. In addition, franchise and public service of public infrastructure are two characteristics in PPP project. Accordingly, five indicators are selected to measure the financing efficiency of PPP projects in this study: the ratio of capital funds, the duration, and the franchise as input indexes; social influence and the turnover ratio of total capital as output indexes.

Five indicators refer to five aspects of financing efficiency in PPP projects. Their names, definitions, notations and formulas for indicators are listed in Table 1.

For making up an effective decision-making unit set, the number of decision-making units should be no less than twice the number of a sum of the input and output indicators; otherwise its ability to distinguish efficiency will decline [7]. So, when measures the relative efficiency of a PPP project, the data of the other nine projects is needed to compose the decision-making unit set. Then input all indexes into DEA model and results can be generated. The calculating model is as Eq. (2) [8].

Table 1. The definition, notation, and calculation formula of five indexes in DEA model

Index	Notation	Meaning	Formulation
Ratio of capital funds	X_1	Percentage of capital amount contributed by partners in the total investment	$X_1 = \frac{\text{capital invested by partners}}{\text{total construction cost}}$
Duration	X_2	Construction time takes in PPP projects, which related to the value of money	$X_2 =$ construction time (month)
Franchise rights	X_3	Privilege the franchise of public infrastructure to co-enterprises	$x_3 = \sum_{t=1}^{m} \frac{B_t}{(1+i^t)}$
Social influence	Y_1	Analyze from the number of service users (μ_1) and related webpage (μ_2)	$Y_{1j} = \frac{\mu_{1j}}{\max\limits_{1 \leq j \leq n} \mu_{1j}} + \frac{\mu_{2j}}{\max\limits_{1 \leq j \leq n} \mu_{2j}}$
Turnover ratio of total capital	Y_2	Evaluate management quality and utilization efficiency of total capital	$Y_2 = \frac{\text{net operating income}}{\text{average total asset}}$

B_t: excess earnings in $t^{th}y$; m : duration of franchise; i: discount rates.

$$\min \left[\theta - \varepsilon\left(\sum_{k=1}^{K} s^{-} + \sum_{l=1}^{L} s^{+}\right)\right]$$
$$\text{s. t.} \begin{cases} \sum\limits_{j=1}^{n} X_{kj}\lambda_j + s^{-} = \theta x_{ko}, & k = 1, 2, 3 \\ \sum\limits_{j=1}^{n} Y_{lj}\lambda_j - s^{+} = y_{ko}, & l = 1, 2 \\ \sum\limits_{j=1}^{n} \lambda_j = 1 & \\ s^{-}, s^{+}, \lambda_j \geq 0, & j = 1, 2, \cdots, n. \end{cases} \quad (2)$$

Three indicators about efficiency can be obtained: technical efficiency θ, scale efficiency θ_{SE} and pure technical efficiency θ_{PE}. θ_{SE} and θ_{PE} separately show the financing efficiency of PPP projects in aspects of scale and pure technique. Only when θ_{PE} and θ_{SE} equal 1 at the same time, it means that the targeted project is in a relatively efficient level among other projects in this decision-making units set. A particular relationship among the three indictors is shown in Eq. (3).

$$\theta = \theta_{PE} \times \theta_{SE}. \quad (3)$$

In addition, the input redundancy s^{+} and the output deficiency s^{-} stem from DEA model indicate targets of every independent indicator for reaching a relative efficiency state. They fail to count uncertainties embodied in PPP projects, while are important raw data for MC simulation.

4.2 Monte Carlo for Simulating Uncertainties

The Integration of MC and DEA is able to build an avenue for simultaneously fulfilling the assessment and prediction of financing efficiency in PPP. Although DEA model have generated an optimal target for each parameter, it is presented

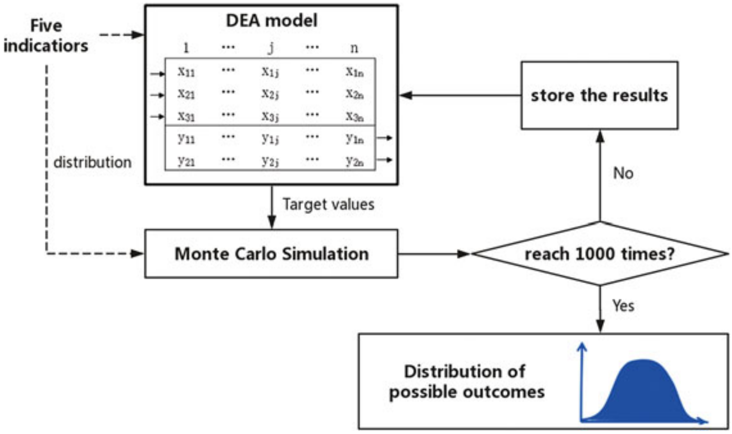


Fig. 1. The flowchart of simulation algorithm integrates Monte Carlo with DEA

as theoretical and fixed data. During the practical development and construction of a PPP project, all the changes in policy, economics, natural environment and project itself generate flexibilities and fluctuations. They may result in a deviation from the expected optimal scheme. MC analysis is strong at simulating realistic uncertainties by randomly sampling the possible parameters set based on their fitted distribution.

The simulation process is the vital part in Monte Carlo analysis. It contains defining distributions of independent parameter, inputting random sampling data set, and gathering results to generate a fitted distribution of dependent indicators. When integrates with DEA model, the sampled random date set need to be put in DEA model for running and giving the results of efficiency indicators, provided that the simulation times does not reach 1000 yet. An aggregation of all the results obtained from every simulation forms a possible financing efficiency distribution. Consequently, uncertainty is considered in the majorization. The flowchart of simulation process of the new model is shown in Fig. 1.

The innovative algorithm combines data processes of MC and DEA, which contribute to a complement of both methods. DEA creates raw data and algorithm for simulation in MC, and MC simulation bridge the gap between reality and theoretical values given by DEA. With the integrated algorithm, a probability distribution of financing efficiency with potential changes in implement will be drawn. It offers a quantitative reference for managers of PPP projects to analyze the feasibility and risk of the optimum proposal, and then make more evidence-based and financing efficient decision.

5 Case Study

5.1 Project Introduction

Australia Adelaide Water Utility project was developed to mitigate the water threat. PPP mode was applied to ally the public and private sector for supporting more efficient and quality service. The United Water Corporation (private sector) cooperated with the Water Company in South Australia (public sector) to construct the PPP project. The overall budget was 4.3 billion included 1.6 billion capital investments, which was raised from both main stakeholders and the others investment from banks and other social capital. During the negotiation stage, the government committed a 27 year franchise right. The completed water utility was expected to serve 5 million citizens and get 521 million's annual operating incomes.

5.2 Simulation and Optimization Analysis

In order to composing an effective decision-making unit, set for DEA analysis, we collected the related data of nine other PPP projects include Bird Nest and the Sydney Olympic Games Stadium. In order to guarantee their comparability, these projects shared similar scale and age of building; further, all the data is transform into the form of ratio of capital funds, duration, franchise rights, social influence and turnover ratio of total capital. The results of relative financing efficiency: pure technical efficiency θ_{PE}, scale efficiency θ_{SE} and technical efficiency θ are calculated with the application of DEA as listed in Table 2.

From the Table 2, a conclusion is drawn that, the financing efficiency is far away from relatively efficient state compared with the other projects in this decision-making unit set. When other projects almost reached 1, a relative efficient state, the pure technical efficiency, scale efficiency and technical efficiency of financing for Australia Adelaide Water Utility are separate to be 0.667, 0.454 and 0.303. Thus, an optimal scheme is in an urge need to improve the situation of relative inefficiency in Australia Adelaide Water Utility PPP project.

DEA model delivers an improvement strategy for relative inefficient units to reach an ideal level, namely a relative efficient financing efficiency state. For Australia Adelaide Water Utility, the initial input and final target given by DEA model are shown in Table 3.

Table 2. Financing efficiency of each project stem from DEA model

Project name	θ_{PE}	θ_{SE}	θ
Australia Adelaide Water Utility	0.667	0.454 irs	0.303
Bird Nest	1	1	1
Sidney Olympic Stadium	1	0.969 irs	0.969
Beijing no. 4 subway	1	1	1
South Africa Ekurhuleni high speed railway	1	0.894 irs	0.894
Perpignan-Figueras rail link	1	1	1
Sidney Harbour Tunnel	1	1	1
Kamchay Hydropower Station	0.725	0.993 irs	0.721
QingDao Sewage Treatment Works	1	1	1
Hong Kong Disneyland	1	1	1

Table 3. Comparison of initial and target data of Australia Adelaide Water Utility

Item	Ratioof capital funds	Duration	Franchise rights	Social influence	Turnover ratio of total capital
Initial input	0.37	36	77.28	0.18	0.12
Target value	0.21	24	11.5	0.27	0.3

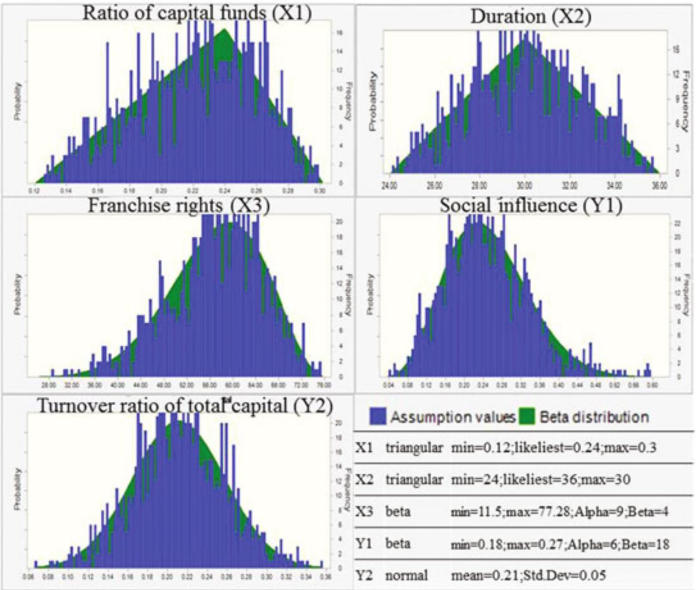


Fig. 2. The flowchart of simulation algorithm integrates Monte Carlo with DEA

The target values are more theoretical, rather than feasible and practical. For in the actual construction, it is surrounded by a changing environment all the time, any uncertainty can result in a deviation from the optimal target. If we predict and simulate all the possibility value of independent parameters, the possible value of dependent parameters can be obtained. MC is created to imitate this process. Ratio of capital funds, duration, franchise rights, social influence and turnover ratio of total capital are five independent parameters in this algorithm, we collect data from over hundred PPP projects so as to generate a fitting probability distribution for each parameter, and they are hypothesized as show in Fig. 2.

After sampling 1000 times in Monte Carlo simulation and running extracted data in DEA model, the crossover process finally generates the probability distribution of pure technical efficiency θ_{PE}, scale efficiency θ_{SE} and technical efficiency θ (Fig. 3). It shows the fitted probability curve for scale efficiency and technical efficiency are lognormal distribution; while for pure technical efficiency

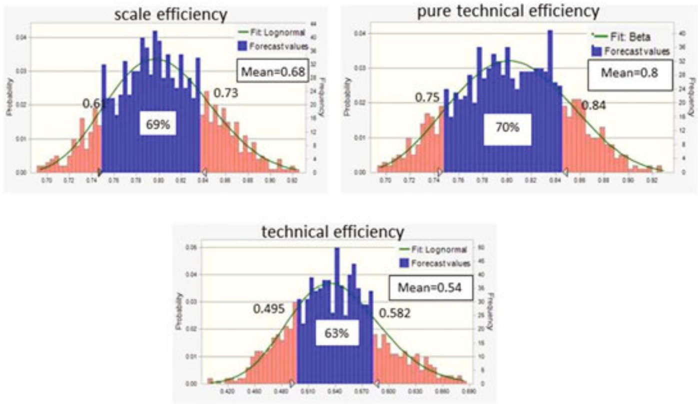


Fig. 3. The flowchart of simulation algorithm integrates Monte Carlo with DEA

is beta distribution. With setting the optimal scheme as a target, the expected value for the pure technical efficiency reaches 0.8, has a promotion over 20% than the initial value; for the scale efficiency reaches 0.68, get a considerable 50% improvement; for the technical efficiency meets 0.54, improves 80%. The blue area highlighted in graphs reveals the most probable range of financing efficiency value with a boundary set at the half of the maximum value.

Finally, we can make a conclusion that the practical implementation of an optimal scheme hardly fulfills the relative efficient state in the end, because of all the uncertainties and change of environment related with PPP projects. While the innovative algorithm proposed in this research offers an approach to predict possible outcomes and distribution of financing efficiency. It enables stakeholders make a better decision among financing strategies and prepare for probable consequences.

6 Conclusions

In this work, we propose an efficient and practical model for estimating and predicting financing efficiency in PPP projects. It integrates DEA and Monte Carlo approach to measure financing efficiency scientifically and objectively, as well as take the uncertainties embodied in PPP project into consider. The innovative model designs a crossover running between DEA and MC, which enable them to compensate for each other's weaknesses. DEA is an effective approach to measure the relative efficiency and delivered an optimal target to reach the relative efficiency state [19]. Wanke, Peter F also applied DEA to analyze scale efficiency of PPP projects in Brazilian ports [16]. While PPP projects are full of risks and uncertainties which will result in a fluctuation from the optimal target, the intrinsic restrictions of DEA fail to take it into consider. A number of scholars, like DijunTan and Yuzba Bahadr, use the features of Monte Carlo to quantify and visualize volatility and flexibility [3,18]. In this manuscript, the MC method is combined with DEA to handle uncertainties through a series of simulation. The proposed algorithm offers a comprehensive method to assess and forecast the financing efficiency in PPP projects.

In the case study of Australia Adelaide Water Utility project, one- thousand-time simulation forms the possible probability distribution of final pure technical efficiency, scale efficiency and technical efficiency of financing. It reveals the possible outcome of financing efficiency in this PPP project when set the optimal value derived from DEA as the target. It shows that, although uncertainties make the relative efficiency hardly to meet a perfectly financing efficient state, while the three efficiency parameters promoted 20%–80% than the initial financing scheme. This result enables the stakeholders to have a full preparation for the probable consequence.

The application of the new method which integrates MC and DEA is not limited in the financing efficiency of PPP projects. It can be extended to various fields whose assessment and simulation process involve multiple indicators and uncertainties. After making corresponding changes in input and output indexes and fit their possibility distribution, the innovative algorithm can be widely used in different areas.

7 Summary of Research Results and Future Study

This study proposed a new model to predict the financing efficiency with fluctuation caused by uncertainties in PPP projects, based on integration of DEA and MC. The case study of Australia Adelaide Water Utility project shows that this innovative method is capable to evaluate and predict flexibilities quantitatively and visually. It also indicates that the uncertainties in PPP projects make the target values in optimal scheme hard to reach, while in this specific project, it also realizes a considerable promotion in all three efficiency indicators.

Further research, such as a more comprehensive indexes system of DEA and subdivide uncertainties in to different categories, is necessary to delve more in-depth and practical in this topic. A more comprehensive and thorough investigation of related factors in DEA method will contribute to a more accurate outcome. Considering that the implementation of PPPs would be affected by economic, social and environmental conditions, sorted uncertainties can facilitate a more accurate calculation. Alternatively, a series of diversified case study need to be carried out in the future to verify and enrich the reliability and feasibility of this new model.

Acknowledgement. The authors would like to thank the National Natural Science Foundation of China for financially supporting this research (Grant No.: 71502011). It is also supported by the Fundamental Funds for Humanities and Social Sciences of Beijing Jiaotong University (Grant No.: 2015jbwj013).

References

1. Bank AD (2013) Public-private partnership operational plan 2012-2020: realizing the vision for strategy 2020: the transformational role of public-private partnerships in Asian development bank operations
2. Broadbent J, Gill J, Laughlin R (2008) Identifying and controlling risk: the problem of uncertainty in the private finance initiative in the UK's national health service. Crit Perspect Account 19(1):40–78
3. Dijun T, Yixiang T (2007) Volatility modeling with conditional volatility and realized volatility. In: International Conference on Management Science and Engineering Management, pp 264–272
4. Dubi A (2000) Monte carlo applications in systems engineering. In: IEEE R&M symposium on a Dubi stochastic modeling of realistic systems topics in reliability & maintainability & statistics tutorial notes
5. Franke M, John F (2011) What comes next after recession? - airline industry scenarios and potential end games. J Air Transp Manage 17(1):19–26
6. Gatti S (2013) Project finance in theory and practice: designing, structuring, and financing private and public projects. Academic Press, Cambridge
7. Li G, Lei Q, Yang Y (2014) The efficiency evaluation of Chinese film and television industry listed companies based on DEA method. In: 2014 Proceedings of the eighth international conference on management science and engineering management. Springer, Heidelberg, pp 83–95
8. Lin Q (2013) Introduction of decision analysis. Tsinghua University Press, Beijing, pp 45–47
9. Liu J, Yang Y (2014) Efficiency evaluation of Chinese press and publication listed companies based on DEA model
10. Martins J, Rui CM, Cruz CO (2014) Maximizing the value for money of PPP arrangements through flexibility: an application to airports. J Air Transp Manage 39:72–80
11. Platon V, Constantinescu A (2014) Monte carlo method in risk analysis for investment projects. Procedia Econ Finance 15:393–400
12. Sastoque LM, Arboleda CA, Ponz JL (2016) A proposal for risk allocation in social infrastructure projects applying PPP in Colombia. Procedia Eng 145:1354–1361
13. Schøyen H, Odeck J (2013) The technical efficiency of Norwegian container ports: a comparison to some Nordic and UK container ports using data envelopment analysis (DEA). Marit Econ Logist 15(2):197–221
14. Verweij S (2015) Producing satisfactory outcomes in the implementation phase of PPP infrastructure projects: a fuzzy set qualitative comparative analysis of 27 road constructions in the Netherlands. Int J Project Manage 33(8):1877–1887
15. Wang H, Huang J, Li H (2017) Local government debt risk, fiscal expenditure efficiency and economic growth. Springer, Singapore
16. Wanke PF, Barros CP (2015) Public-private partnerships and scale efficiency in Brazilian ports: evidence from two-stage dea analysis. Socio-Econ Plann Sci 51:13–22
17. Yu Y, Gu X, Chen Y (2017) Research on the technology transfer efficiency evaluation in industry-university-research institution collaborative innovation and its affecting factors based on the two-stage DEA model. In: 2017 Proceedings of the tenth international conference on management science and engineering management. Springer, Singapore, pp 237–249

18. Yüzbaşı B, Ahmed SE (2015) Shrinkage ridge regression estimators in high-dimensional linear models. Springer, Heidelberg
19. Zhang S, Chan APC et al (2016) Critical review on PPP research - a search from the Chinese and international journals. Int J Project Manage 34(4):597–612
20. Zhongji X (1985) Monte Carlo Approach. Shanghai Science and Technology Press, Shanghai
21. Zhuang M (2012) Research on fiscal control in public-private partnerships of city infrastructure. In: Electronic system-integration technology conference, pp 1–4

The Moderating Effects of Capacity Utilization on the Relationship Between Capacity Changes and Asymmetric Labor Costs Behavior

Abdulellah Azeez Karrar[1,2(✉)], DongPing Han[1], and Sobakinova Donata[1]

[1] School of Management, Harbin Institute of Technology, Harbin 150001, People's Republic of China
[2] Collage of Economics and Administrative, University of Kufa, Najaf 54001, Iraq
karara.alkhaldy@uokufa.edu.iq

Abstract. The study presents an interactive and theoretical research model based on the economic theory of asymmetry behavior, which means labor costs are more likely to increase when activity increases greater than activity decreases. This paper aims to analyze the effect of capacity utilization on prior capacity changes and labor costs behavior. Study methodology tests the hypotheses of nonlinear relationships among variables in the proposed model using multiple regression analysis. Data have been collected from five factories of Iraqi cement firms by physical output data from 2006 to 2015 monthly. Findings reveal that degree of cost asymmetry of manufacturing, selling and administrative, and total labor is associated with capacity changes and that a moderating effect of capacity utilization strongly reveals the opposite pattern supported. Our results from Iraqi manufacturing provide a new evidence on labor costs behavior and contain important implications for both accounting and economic theories. The value of the paper is one of the earliest studies that empirically examine and explain the moderating effects of capacity utilization on employment capacity and labor costs behavior.

Keywords: Capacity utilization · Employment capacity · Labor costs behavior · Asymmetry theory

1 Introduction

Alternative behavior model is an important issue based on deliberate managerial decisions for cost accounting [1]. Underlying the traditional behavior model is based on the activity level that associates changes linearly and proportionately [30]. Many studies have documented an asymmetry behavior between costs and resources in various perspectives [15,22–25,36]. The variances between the level of costs to activity rise changes and the level of costs to activity fall changes called "asymmetry" as sticky, anti-stick costs [7]. Conversely, The asymmetric cost behavior is investigating the consequence cost stickiness on actual and forecast earnings [28].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_21

Managers tend to save capacity when capacity is strained and demand falls then, add capacity when demand grows [7,13,16]. As optimistic expectations lead to saving idle capacity in the sales response to necessary declines but these predictions will make a sticky cost behavior [17]. The asymmetric slack maintenance does not automatically lead to cost stickiness, but if managers retained in response to a previous sales decrease, the significant is stickiness on a prior sales increase and anti-stickiness on a prior sales decrease [11]. These motivations lead us to question is "how capacity utilization moderate the relationship between labor costs and capacity changes?" Capacity utilization is a measurement of the percentage of firm productivity capacity that determine unused resources to total resources [31].

Unemployment fluctuations may make stickiness labor costs in structural parameters because the statistics of unemployment are replicated as well in sticky labor behavior [14]. Empirical implications created a counter cyclical wedge between the real labor and marginal rate [19]. Moreover, the significance of labor is not flexible or inflexible because it is necessary to manage cost structure [6]. Moderating model has certain guiding significance for the application of big data to management decision [37].

Prior studies identified that capacity changes asymmetrically impact on manufacturing labor, selling and administrative, and total costs [5,13]. However, Banker et al. [9] examined the moderating effect of prior periods on the relationship between sales changes and costs behavior. Yet, no research available in literature presents the impact of capacity utilization on asymmetric labor costs behavior.

This paper aims to analyze the impact of capacity utilization on the asymmetry behavior between labor costs and employment capacity, and to explain scientifically the differences without and within the interaction effects.

2 Related Studies and Hypothesis Development

Resources adjustment costs create asymmetry behavior in marginal costs because managerial expectations depend on optimism and pessimism about sales future [21]. Sometimes, uncompleted cost information generates asymmetric results [34]. It depends on how managers use resources and understand their mechanism [33]. Managerial optimism weakness on previous changes called cost sticky, and managerial pessimism weakness on previous changes called anti-sticky [2]. Asymmetry behavior is managerial practical because managers are expected reluctant to lower committed resources in periods of cost-effective growth [8,12]. Limitation of labor costs associated with operating's during sales changes [38]. In addition, high product competition has affected costs behavior that generates Stickiness costs or likely cost responsive [17]. This behavior of labor cost will generate a good control in future because the cost should match with managerial expectations that mean little variations [11]. Overall, The main reason of asymmetry behavior is cost response cannot be a rise or decline fast enough to combine changes in resources [4,30]. Few studies have considered the effect of asymmetric information on reporting decisions and analyze after demand realization under a decentralized setting [27].

The capacity problems and asymmetry behavior positively are associated with control on economics variables such as employees' intensity, production unit, workforce level [15,26,35,39], and machine level. There are behavioral differences between costs categories [24]. Dalla Via and Perego [18] indicated significant labor cost stickiness conditional on a prior sales changes for Italian companies listed. Chen et al. [15] studied the interaction effects among capacity, economics variables, sales changes, and selling and administrative costs. Their results indicate asymmetry behavior and find significant. Azeez et al. [5] examined the relationship of labor costs and employment capacity in which they activate, and discuss how they share to the growth and improvement rate of categories of labor cost with capacity. Banker et al. [11] used employment protection legislation to examine adjustment costs and employment resources in different countries. Their finding showed that costs respond to sales increase greater than sales decrease.

We summarize these predictions by three hypotheses as follows:

Hypothesis 1a (H1a): Conditional on a prior capacity utilization increases, change in manufacturing labor costs to capacity increase is greater than capacity decrease.

Hypothesis 1b (H1b): Moderating effects on a prior capacity utilization increases, change in manufacturing labor costs to capacity increase is less than capacity decrease.

Hypothesis 2a (H2a): Conditional on a prior capacity utilization increases, change in selling and administrative labor costs to capacity increase is greater than capacity decrease.

Hypothesis 2b (H2b): Moderating effects on a prior capacity utilization increases, change in selling and administrative labor costs to capacity increase is less than capacity decrease.

Hypothesis 3a (H3a): Conditional on a prior capacity utilization increases, change in total labor costs to capacity increase is greater than capacity decrease.

Hypothesis 3b (H3b): Moderating effects on a prior capacity utilization increases, change in total labor costs to capacity increase is less than capacity decrease.

3 Research Methodology

3.1 Sample Description

We have empirically investigated 600 practical observation from five factories across hundred twenty months each one from 2006 to 2015 of cement in national Iraqi company. We calculated all log-change using same months in years as indexes of study variables using OLS estimation for regression and the second stage we applied hierarchal moderation regression as function of results.

3.2 Model Development

Our research design is based on the established asymmetric labor costs model of Anderson et al. [1], and Azeez et al. [5], Banker et al. [9] as in equations:

$$\begin{aligned}\ln\left(\frac{MLC_{i,t}-MLC_{i,t-1}}{MLC_{i,t}}\right) = {} & \varphi_0 + \varphi_1 \ln\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right) \\ & + \varphi_2 \mathrm{DEC}_{i,t} \ln\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right) \\ & + \varphi_3 \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right) \\ & + \varphi_4 \mathrm{DEC}_{i,t} \ln\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right) \\ & \times \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right) + \varepsilon_{i,t},\end{aligned} \tag{1}$$

where $\mathrm{MLC}_{i,t}$ is a manufacturing labor cost for firm i at time t and is a non-linear function of independent variables ($\varphi_1, \varphi_2, \varphi_3$, and φ_4). ME is a manufacturing Capacity of employees by hours. $\mathrm{DEC}_{i,t}$ is an indicator set to 1 if $\mathrm{MEq}_{i,t} < \mathrm{MEq}_{i,t-1}$, and set to 0 otherwise. φ_0 is a parameter that estimates the asymmetric labor cost unassociated with employment capacity changes. φ_1 is a parameter that estimates the association between manufacturing labor cost and employment capacity changes increase. φ_2 is a Parameter of asymmetry measurement that estimates the association between manufacturing labor cost and employment capacity changes during increasing and decreasing. φ_3 is a moderating variable that estimates the association between labor cost and capacity utilization changes. φ_4 is a critical value that estimates how capacity utilization moderates the association between manufacturing employment capacity and labor costs behavior. $\varepsilon_{i,t}$ is an error term for variability in labor cost change estimation for firm i at time t.

The capacity utilization was computed for labor hour's measures of the unused and total capacity for each of the periods as in Eq. (2) below:

$$qu = 1 - \frac{\text{unused capacity}}{\text{total capacity}}, \tag{2}$$

where qu is employees Capacity Utilization for each activities. Unused capacity is that hours cannot exceed effective capacity. Total capacity is the maximum hours of output designed for the operation and facility other.

$$\begin{aligned}\ln\left(\frac{\mathrm{S\&ALC}_{i,t}-\mathrm{S\&ALC}_{i,t-1}}{\mathrm{S\&ALC}_{i,t}}\right) = {} & \theta_0 + \theta_1 \ln\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right) \\ & + \theta_2 \mathrm{DEC}_{i,t} \ln\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right) \\ & + \theta_3 \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right) \\ & + \theta_4 \mathrm{DEC}_{i,t} \ln\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right) \\ & \times \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right) + \omega_{i,t},\end{aligned} \tag{3}$$

where: $\mathrm{S\&ALC}_i$, is a selling and administrative labor cost for firm i at time t. $SEq_{i,t}$ is a selling and administrative Capacity of employees by hours, and all other variables were defined previously.

$$\begin{aligned}\ln\left(\frac{\mathrm{TLC}_{i,t}-\mathrm{TLC}_{i,t-1}}{\mathrm{TLC}_{i,t}}\right) = {} & \gamma_0 + \gamma_1 \ln\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right)\\ & +\gamma_2 \mathrm{DEC}_{i,t} \ln\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right)\\ & +\gamma_3 \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)\\ & +\gamma_4 \mathrm{DEC}_{i,t} \ln\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right)\\ & \times \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right) + \alpha_{i,t},\end{aligned} \tag{4}$$

where TLC_i, is a total labor cost for firm i at time t. $\mathrm{MEq}_{i,t}$ is a total Capacity of employees by hours include all of the activities, and remainder of variables are defined previously.

To develop the concept of asymmetry labor costs about employment capacity structure (high utilization). The coefficients ϕ_1, and measure the average percentage increase in labor costs for one present increase in employment capacity, whereas the sum of coefficients $(\phi_1 + \phi_2)$, $(\theta_1 + \theta_2)$ and $(\gamma_1 + \gamma_2)$ in three models to measure the average percentage of decrease in labor costs for one present

Table 1. Definition of variables in empirical models

Variable/Calculation	Description
$\ln\left(\frac{\mathrm{MLC}_{i,t}-\mathrm{MLC}_{i,t-1}}{\mathrm{MLC}_{i,t}}\right)$	Log-change in manufacturing labor costs. All labor costs of manufacturing process and its supporting, payments of industrial activity
$\ln\left(\frac{\mathrm{S\&ALC}_{i,t}-\mathrm{S\&ALC}_{i,t-1}}{\mathrm{S\&ALC}_{i,t}}\right)$	Log-change in Selling & administrative labor costs. All labor costs of selling and administrative process, payments of selling, and administrative activity
$\ln\left(\frac{\mathrm{TLC}_{i,t}-\mathrm{TLC}_{i,t-1}}{\mathrm{TLC}_{i,t}}\right)$	Log-change in total labor costs. All labor costs of manufacturing, supporting, selling, and administrative activities
$\ln\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right)$	Log-change in manufacturing employment capacity. Total hours of employees in manufacturing and supporting activity
$\ln\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right)$	Log-change in selling employment capacity. Total hours of employees in selling, general, and administrative activity
$\ln\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right)$	Log-change in total employment capacity. Total hours of employees in the all activities
$\ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	Log-change in capacity utilization. The percentage rate of each kind of actual capacity to design capacity

Manufacturing labor cost is collected from manufacturing, engineering & services, and quality control activities. Selling and administrative labor cost is collected from marketing and administration activities. Practical employees' capacity is collected from a dataset of the planning department. Unused employees' capacity is calculated based on differences between actual and practical employees capacity. Actual employees' capacity is calculated by divided labor costs on payment rates.

decrease in the employment capacity. The labor costs behavior is asymmetry when the average percentage increase in labor costs not equal the average percentage decrease in labor costs, that is, if labor costs decrease less when employment capacity fall than they increase when employment capacity rise [1,36]. Means the labor costs have sticky behavior (ϕ_1, θ_1 and γ_1, and will be positive, ϕ_2, θ_2 and γ_2, will be negative). The empirical hypothesis for sticky behavior means that $\phi_2 < 0$. This finding an empirical test of Hypotheses 1a, 2a, and 3a. In addition, the current study is exciting literature using the interactive effect of capacity utilization on the asymmetric relationship between employment capacity and labor costs behavior. *Hypotheses 1b*, *2b*, and *3b* imply ϕ_2, are negative and ϕ_4, are positive reflecting stickiness and anti-stickiness respectively [9]. The moderation role of capacity utilization implies that labor costs asymmetry will be less than in the case without the interactive effects. The variables definitions are defined in Table 1.

4 Findings and Discussion

In the current study, we estimate three models for each hypothesis of manufacturing labor costs, selling and administrative labor costs, and total labor costs (Tables 2, 3 and 4). For all of those categories, the estimates indicate significance levels and support our hypotheses. The evidence finds labor costs behavior is sticky on a prior capacity decrease and indicate significant anti-sticky behavior on a prior capacity decrease with interactive effects of capacity utilization (<0 and >0 continues to all models).

4.1 Asymmetric Labor Costs Response to Capacity Changes

In the regression analysis, we find a sticky behavior between labor costs and capacity changes for all of the labor costs categories but in different degrees, the estimates indicate that φ_1 is positive, φ_2 is negative and significant (see model (1) to each Tables 1, 2 and 3). The results found manufacturing, selling and administrative labor, total labor costs are sticky on average by magnitudes of prior employment capacity increase and decrease.

Table 2-panel A. presents a model (1), manufacturing labor costs response to capacity increase is statistically greater than capacity decrease ($\varphi 1 \succ \varphi_1+\varphi_2$). The coefficient is negative (-0.56%, $SE. = 0.18$) and significantly different from zero at the 1 % (t-statistics -3.176). On average, the manufacturing labor costs increase 0.74% per 1% increase in employment capacity (φ_1) and costs decrease by 0.18% per 1% decrease in employment capacity ($\varphi_1 + \varphi_2$). The results show that the adjusted R^2 is 0.71, the model is significant ($p < 0.001$) level. Means the behavior between employment capacity change and manufacturing labor cost is sticky (*H1a*). However, Selling and administrative labor costs response to capacity increase statistically is greater than capacity decrease ($\theta_1 > +\ 2$). The coefficient is negative and significant (-0.40%, $SE. = 0.12$, t-statistics -3.25). On average, the Selling and administrative labor costs increase 0.71%

Table 2. Validation test of the sticky behavior: Nonlinear analysis of moderation test among employment capacity, capacity utilization, and manufacturing labor costs change

Panel A: Regression analysis: direct effects model 1 Dependent variable = manufacturing labor cost (MLC)				
Variable	Parameter	Parameter estimate	Standard error	Parameter significance (t-statistics)
Intercept	φ_0	0.15	0.08	0.077
		(?)		−1.77
$\ln\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right)$	φ_1	0.74	0.14	0
	Asymmetric measure	(+)		−5.22
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right)$	φ_2	−0.56	0.18	0.002
		(−)		(−3.176)
$\ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	φ_3	0.31	0.07	0.015
		(+)		−4.432
Adjusted R2		0.706		
F-value		481.03		
Significant level		0		
Panel B: Moderation analysis: interactive effects of capacity utilization - model 2				
Intercept	φ_0	0.19	0.35	0.593
		(?)		−0.53
$\ln\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right)$	φ_1	0.67	0.09	0
	Asymmetric measure	(+)		−7.26
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right)$	φ_2	−0.22	0.34	0.516
		(−)		(−0.65)
$\ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	φ_3	0.71	0.13	0.001
		(+)		−5.359
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{MEq}_{i,t}-\mathrm{MEq}_{i,t-1}}{\mathrm{MEq}_{i,t}}\right) \times \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	φ_4	0.11	0.02	0
		(+)		−6.047
Adjusted R2		0.723		
F-value		391.448		
Significant level		0		

The results present a first examination that explain how the interaction affects of capacity utilization impact on the relation between costs behavior and employment capacity changes. The sample consists of 600 factor-month observations between 2006 and 2015 from Iraqi industry on cement produce. All t-statistics calculate by using significant indicate at the 1%, 5%, 10% level respectively.

Table 3. Validation test of the sticky behavior: Nonlinear analysis of moderation test among employment capacity, capacity utilization, and selling and administrative labor costs change

Variable	Parameter	Parameter estimate	Standard error	Parameter significance (t-statistics)
Panel A: Regression analysis: direct effects model 1 Dependent variable = selling and administrative labor cost (S&ALC)				
Intercept	θ_0	0.13	0.08	0.09
		(?)		−1.7
$\ln\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right)$	θ_1	0.71	0.14	0
	Asymmetric measure	(+)		−5.19
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right)$	θ_2	−0.4	0.12	0.001
		(−)		(−3.25)
$\ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	θ_3	0.16	0.04	0.093
		(+)		−4.3
Adjusted R2		0.704		
F-value		475.992		
Significant level		0		
Panel B: Moderation analysis: interactive effects of capacity utilization C model 2				
Intercept	θ_0	0.4	0.33	0.224
		(?)		−1.22
$\ln\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right)$	θ_1	0.74	0.1	0
	Asymmetric behavior	(+)		−7.11
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right)$	θ_2	−0.19	0.05	0.015
		(−)		(−3.67)
$\ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	θ_3	0.7	0.25	0.042
		(+)		−2.79
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{SEq}_{i,t}-\mathrm{SEq}_{i,t-1}}{\mathrm{SEq}_{i,t}}\right) \times \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	θ_4	0.28	0.04	0
		(+)		−6.81
Adjusted R2		0.725		
F-value		395.823		
Significant level		0		

All t-statistics calculate by using significant indicate at the 1%, 5%, 10% level respectively.

per 1% increase in employment capacity (θ_1) and costs decrease by 0.31% per 1% decrease in employment capacity ($\theta_1 + \theta_2$), for which Selling and administrative labor costs expected to be sticky (*H2a*). The adjusted R2 is 0.70, where F-value is 475.99, the model is significant ($p < 0.001$) level. (Presented in Table 3-panel A.).

Finally, Table 4-panel A. in model (1), presented the empirical examination of the relationship between total labor costs and employment capacity changes. Total labor costs asymmetrically respond to capacity changes. The coefficient is negative and significant (-0.58%, $SE. = 0.08$, t-statistics -7.68). On average, the total labor costs increase by 0.67% per 1% increase in employment capacity (γ_1) and they decrease by 0.09% per 1% decrease in capacity ($\gamma_1 + \gamma_2$). The results show that the adjusted R^2 is 0.09, where F-value is 21.09, the model is significant ($p < 0.001$) level. The result is consistent with (*H3a*).

4.2 Moderation Analysis: Interactive Effects of Capacity Utilization

In extension analysis, we estimate interactive models for each of the main components of labor costs (manufacturing labor, selling and administrative labor, and total labor), for extending literature of asymmetric cost behavior. The estimates indicate significant stickiness conditional on a prior capacity decrease and significant anti-stickiness conditional on a prior capacity decrease with moderating effects of capacity utilization change ($\varphi_2 < 0$ and $\varphi_4 > 0$ respectively) suggesting that capacity utilization changes are related to the effects of employment capacity changes on labor costs behavior. These findings support *H1b*, *H2b*, and *H3b*.

Moderation test was conducted to examine the interacting effect of prior sales changes between sales changes and asymmetric cost behavior [9]. The labor costs exhibit significant stickiness without the interactive effects of capacity utilization ($\varphi_2 = -0.22\%$, $SE. = 0.34$, t-statistics -0.65), ($\theta_2 = -0.19\%$, $SE. = 0.05$, t-statistics -3.67), and ($\gamma_2 = -0.48\%$, $SE. = 0.06$, t-statistics -7.51) respectively, but the labor costs reveal the opposite pattern of significant anti-stickiness within the interactive effects of capacity utilization ($\varphi_4 = 0.11\%$, $SE. = 0.02$, t-statistics 6.05), ($\theta_4 = 0.28\%$, $SE. = 0.04$, t-statistics 6.81), and ($\gamma_4 = -0.13\%$, $SE. = 0.06$, t-statistics 2.16). These results document that labor costs are description of a broader pattern of asymmetric cost behavior, which extends to all the major components of labor costs for physical input quantity (employment capacity) for labor costs behavior.

In additional, Banker et al. [9] indicated that costs respond to activity increase is greater than activity increase within the moderation effects, which estimated parameter of total employees costs increase equal 0.62 stronger than 0.42 within the interactive effects. Whereas our finding indicate that coefficients of asymmetry measure less than within the interactive effects ($\varphi_2 = 0.56, \theta_2 = -0.40, \gamma_2 = -0.58$) $>$ ($\varphi_2 = -0.22, \theta_2 = -0.19, \gamma_2 = -0.48$) respectively and capacity utilization change was observed to strengthen the relationship between capacity changes and labor costs behavior ($\Delta R^2 = 0.017$, $\Delta R^2 = 0.021$, $\Delta R^2 = 0.009$, $p < 0.001$). The results also underscore the value-added of interactive analysis of capacity utilization, which in the moderation models of categories of labor costs are important for drawing accurate implications about the natural of costs behavior. While Banker et al. [9] considered the moderation of two-period analysis is value- added for theory of asymmetric cost behavior.

Table 4. Validation test of the sticky behavior: Nonlinear analysis of moderation test among employment capacity, capacity utilization, and total labor costs change

Variable	Parameter	Parameter estimate	Standard error	Parameter significance (t-statistics)
Panel A: Regression analysis: direct effects C model 1 Dependent variable = total labor cost (TLC)				
Intercept	γ_0	0.66	0.29	0
		(?)	−2.24	
$\ln\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right)$	γ_1	0.67	0.09	0
	Asymmetric behavior	(+)		−7.16
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right)$	γ_2	−0.58	0.08	0
		(−)		(−7.68)
$\ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	γ_3	0.74	0.45	0.1
		(+)		−1.62
Adjusted R2		0.091		
F-value		21.09		
Significant level		0		
Panel B : Moderation analysis: interactive effects of capacity utilization - model 2				
Intercept	γ_0	0.9	0.32	0
		(?)		−2.76
$\ln\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right)$	γ_1	0.86	0.11	0
	Asymmetric behavior	(+)		−7.29
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right)$	γ_2	−0.48	0.06	0
		(−)		(−7.51)
$\ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	γ_3	0.41	0.2	0.049
		(+)		−2.007
$\ln \mathrm{DEC}_{i,t}\left(\frac{\mathrm{TEq}_{i,t}-\mathrm{TEq}_{i,t-1}}{\mathrm{TEq}_{i,t}}\right) \times \ln\left(\frac{qu_{i,t}-qu_{i,t-1}}{qu_{i,t}}\right)$	γ_4	0.13	0.06	0.031
		(+)		−2.16
Adjusted R2		0.1		
F-value		17.084		
Significant level		0		

All t-statistics calculate by using significant indicate at the 1%, 5%, 10% level respectively.

4.3 Discussion

This research was conducted to analyze the moderation effects of capacity utilization changes as they related to asymmetric costs behavior at manufacturing sector firms in Iraq. Until recently, most studies have been conducted to investigate why change in labor costs behavior to activity is asymmetry [3,8,18,20,29], specially Banker et al. [9] conducted on moderating effects but they ignored the important dimension of capacity utilization and also focused on multiple periods analysis for same variables to overcome these issues, this research has presented an interactive model based on data collection from multiple and different variables. Maintenance management is necessary an effective system to display the surface temperature to reduce the capacity losses [32].

From the variable is used in the theoretical research model, three nonlinear relationships were proposed and tested. The results found capacity utilization changes can affect asymmetric cost behavior. Conversely, labor costs and their categories have associated with employment capacity changes in different degrees such as significant stickiness conditional on a prior capacity decrease and significant anti-stickiness conditional on an interaction of capacity decreases with capacity utilization increases. Finally, capacity utilization can decrease the degree of sticky measure. The results from the nonlinear relationships tested to confirm and align with the existing literature [1,3,5,9].

Anderson et al. [1] Documented that behavior between costs and resources changes is asymmetry because costs respond to resources increase greater/less than resources decrease that rejected the traditional models about fixed and variable changes. However, costs behavior is sticky when prior sales increase, whereas costs behavior is anti-sticky when prior sales change decrease. The costs are anti-sticky when capacity utilization is unusually low [9]. On the other hand, Azeez et al. [5] examined the labor costs behavior when employment capacity change is increasing and decreasing for existing asymmetric cost behavior by using physical output data. Finally, the results found capacity utilization is able to moderate the relationship between capacity change and labor costs behavior, and suggest new avenues of exploration for future studies.

5 Conclusion

Significance of the relationship between costs behavior and resources changes has been an extension of the asymmetric cost behavior. Asymmetry phenomenon is a new thinking created by [1] but it still under discussion because they did not use many drivers, their findings used sales revenues change only and ignored a physical output data for two reasons: physical data is not available, and sales revenues typically is a more an appropriate empirical measure for activity than physical data [10]. The current study provides an empirical examining that added a new dimension for explaining how capacity utilization change moderates the relationship between capacity changes and labor costs behavior. In addition, our findings showed a complex fundamental design of asymmetric cost behavior that combines two roles: labor costs behavior is sticky on a prior capacity increase without the interaction of capacity utilization changes. Second, labor costs behavior is anti-sticky on a prior capacity increase within the interaction of capacity utilization changes, means the moderation effects of capacity utilization made the opposite of the standard predictions.

The findings of the current study have revealed that capacity utilization changes significantly have affected labor costs behavior. Furthermore, these predictions explain the influence of managerial expectations for future on firm resources such as the general structure of optimal decisions with resources adjustment costs, which generate asymmetry behavior. Finally, using capacity utilization changes with asymmetric relations can decrease the degree of sticky behavior between labor costs and capacity change, interaction capacity utilization changes with a prior capacity decrease can converse the asymmetry degree from sticky to anti-sticky. The results of moderation are unique and would provide productive avenues for future studies.

Acknowledgements. The authors would like to thank Iraqi Cultural Attaché in China and national company for supporting a research about data collecting and their corresponding.

References

1. Anderson MC, Banker RD, Janakiraman SN (2003) Are selling, general, and administrative costs “sticky”? J Account Res 41(1):47–63
2. Anderson MC, Lee JH, Mashruwala R (2015) Cost stickiness and cost inertia: a two-driver model of cost behavior. Social Science Electronic Publishing SSRN 2599108
3. Anderson SW, Lanen WN (2007) Understanding cost management: what can we learn from the evidence on ‘sticky costs’? SSRN Electron J SSRN 975135
4. Argilés JMA, Garcíablandón J (2007) Cost stickiness revisited: empirical application for farms. Revista Espanola De Financiacion Y Contabilidad 38(187):579–605
5. Azeez KA, DongPing H, Mabula JB et al (2016) Using capacity utilization to measure asymmetric labor cost behavior and capacity expansion. In: Management science and engineering (ICMSE)
6. Babecky J, Caju PD et al (2012) How do European firms adjust their labour costs when nominal wages are rigid? Labour Econ 19(5):792–801
7. Balakrishnan R, Labro E, Soderstrom NS (2014) Cost structure and sticky costs. J Manage Account Res 26(2):91–116
8. Banker RD (2006) Labor market characteristics and cross-country differences in cost stickiness. SSRN Electron J
9. Banker RD (2014) The moderating effect of prior sales changes on asymmetric cost behavior. J Manage Account Res 26:221–242
10. Banker RD, Byzalov D (2014) Asymmetric cost behavior. J Manage Account Res 26(2):43–79
11. Banker RD, Basu S et al (2016) The confounding effect of cost stickiness on conservatism estimates. J Account Econ 61(1):203–220
12. Bugeja M, Lu M, Shan Y (2015) Cost stickiness in australia: characteristics and determinants. Aust Account Rev 25(3):248C261

13. Cannon JN (2014) Determinants of 'sticky costs': an analysis of cost behavior using united states air transportation industry data. Account Rev 89(5):1645
14. Casares M, Moreno A, Vázquez J (2009) Wage stickiness and unemployment fluctuations: an alternative approach. Series 3(3):395–422
15. Chen CX, Hai LU, Sougiannis T (2012) The agency problem, corporate governance, and the asymmetrical behavior of selling, general, and administrative costs. Contemp Account Res 29(1):252C282
16. Chen JV, Kama I, Lehavy R (2015) Management expectations and asymmetric cost behavior. Technical report, Working Paper, University of Illinois at Chicago, Tel Aviv University, and University of Michigan
17. Cohen S (2015) The sticky cost phenomenon at the local government level: empirical evidence from Greece. Social Science Electronic Publishing SSRN 2575530
18. Dalla Via N, Perego P (2015) Sticky cost behaviour: evidence from small and medium sized companies. Account Finance 54(3):753–778
19. Dicecio R (2009) Sticky wages and sectoral labor comovement. J Econ Dyn Control 33(3):538–553
20. Gu Z (2016) The political economy of labor cost behavior: evidence from china. Social Science Electronic Publishing SSRN 2786533
21. Hamermesh DS, Pfann GA (1996) Adjustment costs in factor demand. J Econ Lit 34(3):1264–1292
22. Ibrahim AEA (2015) Economic growth and cost stickiness: evidence from Egypt. J Financ Report Account 13(1):119–140
23. Jang Y (2016) Asymmetric cost behavior and value creation in m&a deals. Social Science Electronic Publishing SSRN 2824132
24. Kama I, Dan W (2010) Do managers' deliberate decisions induce sticky costs? SSRN Electron J 112(1):39–47
25. Kama I, Dan W (2013) Do earnings targets and managerial incentives affect sticky costs? J Account Res 51(1):201C224
26. Kee RC (2003) Operational planning and control with an activity-based costing system. Adv Manage Account 11:59–84
27. Liao Y, Shen W, Lev B (2017) Asymmetric information effect on transshipment reporting strategy
28. Malik M (2012) A review and synthesis of 'cost stickiness' literature. SSRN Electron J SSRN 2276760
29. Mohammad MN, Masoumeh N (2011) Evidence of cost behavior in Iranian firms. In: 2011 international conference on advancements in information technology, vol 20, pp 254–258
30. Nematollahi M (2013) Investigation the relation between costs and revenues in Iranian firms. N Y Sci J 6(11):52–57
31. Nyaoga RB, Wang M, Magutu PO (2015) Does capacity utilization mediate the relationship between operations constraint management and value chain performance of tea processing firms? evidence from Kenya. Int Strateg Manage Rev 3(1–2):81–95
32. Ramirez IS, Munoz CQG, Marquez FPG (2017) A condition monitoring system for blades of wind turbine maintenance management. In: Proceedings of the tenth international conference on management science and engineering management, pp 3–11
33. Teece DJ, Pisano G, Shuen A (2009) Dynamic capabilities and strategic management. Oxford University Press, New York
34. Von Hippel E (1994) "Sticky Information" and the locus of problem solving: implications for innovation. Manage Sci 40(4):429–439

35. Watts T, Mcnair CJ et al (2009) Structural limits of capacity and implications for visibility. J Account Organ Change 5(2):294–312
36. Weidenmier ML, Subramaniam C (2003) Additional evidence on the sticky behavior of costs. Social Science Electronic Publishing
37. Yang A, Lan X et al (2017) An empirical study on the prisoners' dilemma of management decision using big data. Springer, Singapore
38. Zanella F, Oyelere P, Hossain S (2015) Are costs really sticky? evidence from publicly listed companies in the uae. Appl Econ 47(60):6519–6528
39. Zhuang ZY, Chang SC (2015) Deciding product mix based on time-driven activity-based costing by mixed integer programming. J Intell Manuf 4:1–16

The Empirical Analysis of the Impact of Technical Innovation on Manufacturing Upgrading-Based on Subdivision Industry of China

Dan Jiang(✉) and Yuan Yuan

Chongqing Technology and Business School,
Chongqing 600000, People's Republic of China
719221214@qq.com

Abstract. At this stage, changing the mode of economic growth and optimizing the industrial structure is the important content for China's economy restructuring and steady growth, so the adjustment of industrial structure is imminent, especially manufacturing. This paper describes the present situation of China's technology innovation and the development of manufacturing, and analyzes the problems of manufacturing upgrading, and then uses the data of China' 27 manufacturing sub-sectors from 2005 to 2014 and establishes the regression model. The results showed that technical innovation has a positive role in promoting the upgrading of China's manufacturing, but its effect is still affected by certain factors, and the differences on technical innovation are obvious. Finally, in view of the problems found in the research, the paper puts forward some policy suggestions for the upgrading of China's manufacturing industry, such as increasing R&D investment, building system of financial support, and establishing a technical alliance about industrial-academic-research.

Keywords: Technical innovation · Manufacturing upgrading · Labor-intensive industry · Capital-intensive industry · Technology-intensive industry

1 Introduction

The manufacturing industry reflect a country's productivity level, as well as present the national competitiveness. It is the cornerstone of the realization of social progress and national prosperity. With the advantage of reform and opening policy, the labor force, some other natural resources and other factors, China' manufacturing industry has created a miracle and continued growth of economy. In 2010, China's manufacture value added exceeded the United States, becoming a manufacturing country, is worthy of the name. As shown in Fig. 1.

Since the financial crisis in 2008, the development of the real economy has been paid more and more attention by the countries all over the world. Therefore,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_22

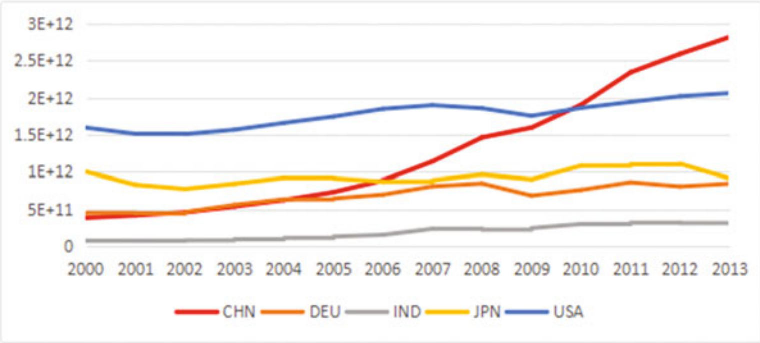


Fig. 1. Comparison of manufacturing value added (dollar)

the development of manufacturing industry and the expansion of exports are the key to the transformation of the world economic structure. The developed countries have taken measures to revitalize the manufacturing sector. For instance, the USA wishes "re industrialization" once again to opened the gap with the developing countries, to further promote its economic and social development, to enhance its competitiveness in the international manufacture. Germany issued "industry 4.0" which was proposed to establish an information physics system network in 2013, making itself the creator and supplier of advanced intelligent manufacturing technology. It provides the opportunities of technological innovation for China's implementation of "made in China 2025". At the same time, some developing countries take advantage of opportunities of the new round of international industrial transfer, promote the industrialization process, thereby exacerbating China's manufacturing competitive environment.

With China's rising labor costs and resource depletion, the advantage of traditional manufacturing relying on cheap labor and resources is gradually lost. In a word, we must rely on technological innovation to speed up the restructuring of China's manufacturing and the transformation of development mode. But how does technological innovation affect the manufacturing upgrading of different factors intensive industrial structure? This is an urgent problem in the process of China's industrialization. In view of this, based on the status of manufacturing upgrading and the thinking of upgrading power, this paper focuses on the effect of technical innovation on manufacturing upgrading.

2 Literature Review

Foreign scholars have carried on the thorough discussion on the technology innovation and the industrial structure evolution, but it mostly based on the industrial level. Guan [4] analyzed the relationship between technological innovation and industrial competitiveness in the ability of R&D personnel's learning and knowledge acquisition, research and development ability, and the research showed that technology innovation is the key way to promote the competitiveness of the industry. Castellacci [2] studied the relationship between the external innovation environment of technology innovation and industrial competitiveness, the results show that the economic policies based on market oriented and institutional arrangements will directly affect the innovation mode, and play the role in industry competitiveness. Cohen [3] considered the Schumpeterian hypotheses relating innovation to market structure and firm size, and considers in more depth the role of firm characteristics and industry-level variablesłbroadly characterized as reflecting demand, technological opportunity, and demand conditionsłin affecting firms' innovative activity and performance. Adak [1] focused on the influence of technological progress and innovation on the Turkish economy, and got the result that a significant effect of technological progress and innovation on economic growth. Hong et al. [5] found that government grants exert a negative influence on innovation efficiency of high-tech industries. However, the impact of private R&D funding is significant and positive. In view of the problem of upgrading the structure of manufacturing industry, this paper will not discuss it in detail.

Domestic scholars focus on the research on the mechanism and path of technological innovation to upgrade the manufacturing industry. Sun and Ye [6] analyzed the role and mechanism of innovation for transformation and upgrading of manufacturing industry based on the connotation of the innovation drive, and considered innovation driven to give impetus to manufacturing from three aspects of power dimension, elements dimensions and competition dimension. Zhao and Qiao [8] studied the upgrading mode of Shanghai' manufacturing industry, and pointed out that Shanghai needed to promote technology improvement and innovation of the traditional manufacturing, and created growth industry through the industrialization of high and new technology industry, and pushed on the dynamic integration of producer services and advanced manufacturing technology, all of which strengthen policy guidance so that it can realize comprehensively manufacturing upgrading.

From what has been discussed above, the existing literature on the analysis of the technology innovation's influence on manufacturing upgrading, which was mostly based on regional classification, and there is less research based on manufacturing industry so that it failed to reveal the impact of technological innovation on Internal structure changes in manufacturing industry. Based on the panel data model, the paper made an empirical analysis of the differences in the impact of scientific and technical personnel, R&D internal expenditure, sales revenue of new products, the number of effective invention patent, cost of technical renovation, main business profit on the manufacturing upgrading of labor-intensive industry, capital-intensive industry and technology-intensive industry. Through the analysis of the technology innovation's influence on different kinds of different factors intensive industries, that can provide guidance for our country to make the corresponding different incentive policy on technological innovation, and avoid invalid policy, and make the technological innovation to better play its role so to realize the upgrading of the manufacturing structure. It has certain practical significance.

3 Theoretical Basis

According to the classification standards of the National Bureau of Statistics, the manufacturing industry is divided into 31 categories, due to the lack of data during some years, this paper chooses 27 Subdivision manufacturing industry of national above-scale industrial as the research object. Referring to the classification method of the factor intensity and Wang and Dong' classification method [7] based on the structure of the manufacturing, the 27 Subdivision industry of China's manufacturing industry is divided into labor-intensive industry, capital intensive industry and technology intensive industry. The specific classification is shown in Table 1:

Table 1. Classification of different factors intensive industries

Manufacturing industry type	Industrial distribution
Labor-intensive industry	Agricultural food processing industry, food manufacturing, beverage and refined tea industry, tobacco industry, textile industry, textile and apparel and footwear products industry, leather and feather industry, Wood processing industry
Capital-intensive industry	Furniture manufacturing, paper making and paper products, printing and recording media copying, Stationery industry, Oil processing, coking & nuclear fuel processing, Chemical Fiber Industry, Rubber and plastic products industry, non-mental mineral product industry, Ferrous metal smelting and rolling processing industry, Non-ferrous metal smelting and rolling processing industry, Metal products manufacturing
Technology-intensive industry	Pharmaceutical Industry, the manufactures of chemical materials and products, General Equipment Manufacturing sector, Special Equipment Manufacturing Sector, Electronic and Communication Equipment Manufacturing Sector, Electrical machine equipment manufacturing, Transportation Manufacturing, instruments and cultural office machinery

4 Empirical Analysis

4.1 Index Selection

(1) Manufacturing Upgrading Index Selection
In existing research, the value added rate of manufacturing industry is generally considered as a comprehensive reflection of the economic benefits and growth quality of the manufacturing industry. There are also some scholars having used industrial profit and tax amount, but the main business profit of the Subdivision manufacturing industry is chosen as an index to measure the manufacturing upgrading. The reason why the author choose main business profit is that manufacturing upgrading means improving the competitiveness and profitability of enterprises. On the other hand, choosing the main business profit instead of net profit ratio, that can eliminate government subsidies which enterprises receive from government and other non-recurring profit, at the same time, taking into account the availability of data, this paper chooses main business profit as the variable of manufacturing upgrading.

(2) The Index Selection of Technological Innovation
In the measurement of technological innovation by domestic scholars, scientific and technological personnel, R&D investment intensity and other indicators are generally accepted evaluation indicators. However, considering the transformation and application of technological innovation achievements, this paper chooses scientific and technical personnel, R&D internal expenditure, sales revenue of new products, the number of effective invention patent, cost of technical renovation to measure the level of technology innovation according to the availability of data. The author tend to specifically study the impact of technological innovation on the manufacturing upgrading, and don't just use a single indicator of technological innovation.

(3) The Selection of Control Variable
In this paper, adding the input of labor and capital which his regarded as control variables into the model, that can reflect the impact of technical innovation and factor input on the changes of manufacturing added value. The net value of fixed assets of industrial enterprises above designated size approximately represent material capital investment, and annual average number of employed persons in all sectors of the manufacturing industry means the number of workers in the manufacturing.

4.2 Data Description

The data which is used in the paper come from the "Statistical yearbook in China", "Industrial Statistics Yearbook in China", "Labor Statistics Yearbook in China", "Statistical yearbook of scientific and technological activities of industrial enterprises", Some of the data are compiled according to the original data of the statistical yearbook.

(1) Comparison of Scientific and Technical Personnel of Various Types of Manufacturing Industry in China
As shown in Fig. 2, all kinds of scientific and technical personnel of manufacturing industry of our country from 2005 to 2008 has maintained steady growth, due to the financial crisis in 2008, from 2008 to 2009 it showed a downward trend, various types of science and technology personnel began to increase slowly after 2009, at the same time, China's scientific and technical personnel of labor-intensive industry and capital -intensive industry has been lagging behind in technology-intensive industry and the total is about half.

(2) Comparison of R&D Internal Expenditure of Various Types of Manufacturing Industry in China
As shown in Fig. 3, General speaking, among 2005–2014, China's various types of R&D internal expenditure of manufacturing industry is up forwards except among 2008–2009. However, R&D internal expenditure of labor-intensive industry and capital-intensive industry is slower than the growth of technology-intensive industry.

(3) Comparison of Sales Revenue of New Products of Various Types of Manufacturing Industry in China
As shown in Fig. 4, China's sales revenue of new products of technology-intensive industry grew so slowly and it almost unchanged. It shows that China's technology-intensive industry has poor ability of technology transformation, and the scientific and technological achievements can not be effectively translated into practical productive forces. But the capital intensive manufacturing and labor-intensive manufacturing industry has maintained a steady growth. Though the number of effective invention patent of capital and labor intensive manufacturing is far less than the technology intensive industry, the latter continuing innovation investment will make the sales revenue of new products greatly improved.

(4) Comparison of Effective Invention Patent Number of Various Types of Manufacturing Industry in China
As shown in Fig. 5, the number of effective invention patent in China's technology-intensive industry went through a growth spurt, and those of labor-intensive industry and capital-intensive industry increased much slower. Perhaps China's technology-intensive industry more and more pay attention to innovation output and the other two industries funding for innovation is not enough.

(5) Comparison of Cost of Technical Renovation of Various Types of Manufacturing Industry in China
As shown in Fig. 6, cost of technical renovation in China's labor-intensive industry grew slowly, until the peak reached early 2013, after that it was a downward trend. While cost of technical renovation in China's technology-intensive industry basically is on the rise from 2005 to 2010, and exceeded the cost of technical renovation capital- intensive industry in 2010, then it began to decline in 2011, finally it reached at the same level as capital-intensive industry in 2014.

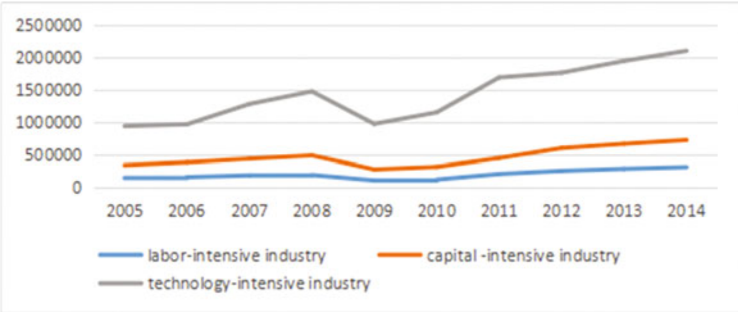


Fig. 2. Comparison of scientific and technical personnel

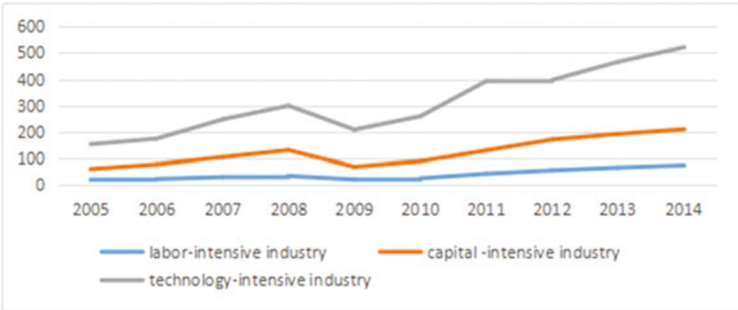


Fig. 3. Comparison of R&D internal expenditure (billion)

(6) Comparison of Main Business Profit of Various Types of Manufacturing Industry in China
As shown in Fig. 7, main business profit in the three kinds of China's manufacturing overall were rising, and those of technology-intensive manufacturing is much more profitable than capital-intensive industry and labor-intensive industry, but capital and labor intensive manufacturing industry's main business profits are similar.

4.3 The Construction of the Model

According to the selected model variables, based on the augmented Cobb Douglas production function, this paper establishes econometric model as follows:

$$Y_{it} = e^{c+\delta_1 \ln RD_{it}+\delta_2 \ln ST_{it}+\delta_3 \ln EP_{it}+\delta_4 \ln PS_{it}+\delta_5 \ln TR_{it}+u_{it}} K_{it}^{\alpha} L_{it}^{\beta}. \quad (1)$$

Among them, Y_{it} represents the main business profit of manufacturing during t period i industry; K_{it}, L_{it} represents fixed capital investment and the amount of labor during t period i industry respectively; RD_{it}, ST_{it}, EP_{it}, PS_{it} and

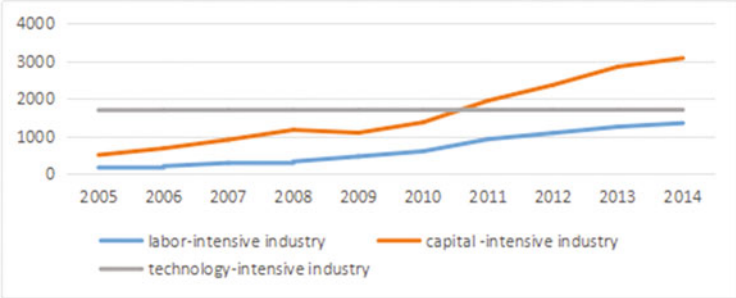


Fig. 4. Comparison of sales revenue of new products (billion)

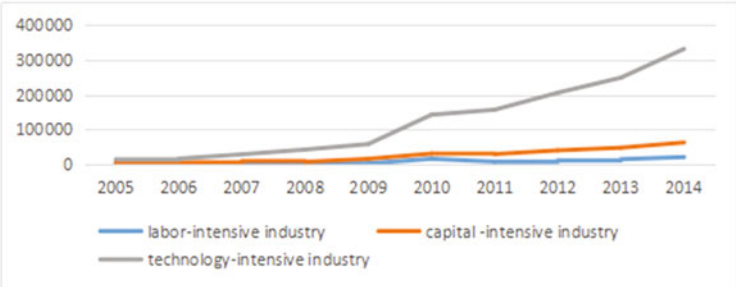


Fig. 5. The number of comparison of effective invention patent

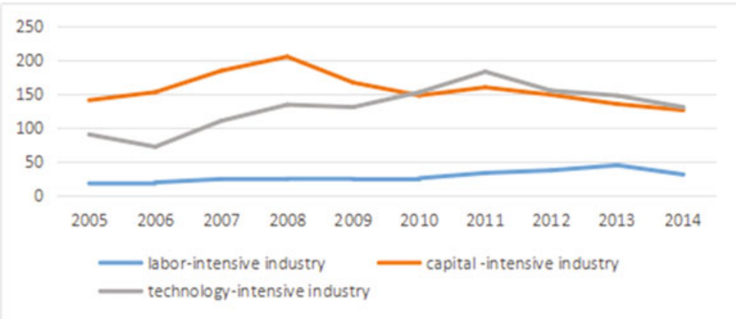


Fig. 6. Comparison of cost of technical renovation (billion)

TR_{it} represents R&D internal expenditure, scientific and technical personnel, the number of effective invention patent, sales revenue of new products, cost of technical renovation during t period i industry, u_{it} represents the random error.

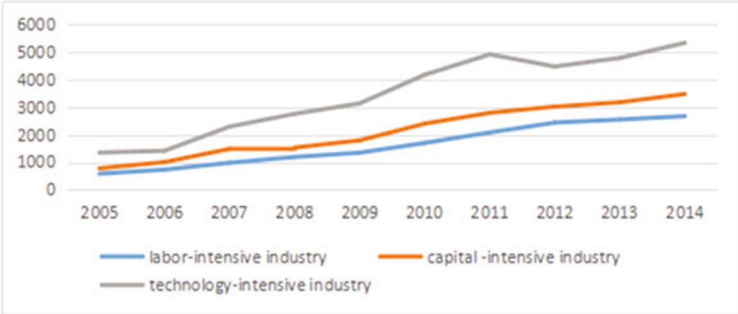


Fig. 7. Comparison of main business profit (billion)

Before testing the data of model, the Eq. (1) on both sides are processed into logarithmic for linear regression model, the model as follows:

$$\begin{aligned} \ln Y_{it} = & \; c + \alpha \ln K_{it} + \beta \ln L_{it} + \delta_1 \ln RD_{it} + \delta_2 \ln ST_{it} + \delta_3 \ln EP_{it} + \delta_4 \ln PS_{it} \\ & + \delta_5 \ln TR_{it} + u_{it}. \end{aligned}$$

4.4 Variable Inspection

(1) Correlation Analysis of Variables
The following is the correlation analysis of selected variables, and from Table 2 we can see that independent variables were significantly correlated with the dependent variable. Specifically, the stronger technique innovation capability is, the higher the main business profit of the industry is. Besides, the capital and labor input is in positive correlation with the operating profits. At the same time, the respective variables are also highly relevant. The explanatory variables including scientific and technical personnel, R&D internal expenditure, sales revenue of new products, the numbers of effective invention patents and cost of technical renovation was significantly correlated in 1% level, and the correlation coefficient is greater than 0.8. It reveals that these indicators selection of technical innovation is over-repeated, and the correlation is very strong.

(2) Empirical Results
After multicollinearity inspection of several variables, we found that VIF is 20.72 in average, more than 10, which shows that there is serious multicollinearity between variables. Via detecting x_3 and x_4 linear with other explanatory variables exist serious, so they should be removed in the model. The value of VIF is less than 10 after deletion. So the final determination of the independent variable is x_1, x_2, x_5 and x_6 in this paper.

Because different manufacturing industries with individual differences and mixed OLS cannot reflect individual differences, it is imperative to use fixed effects and random effects model for regression. Considering the upgrading of industrial structure with inertia index, this paper uses dynamic panel model

Table 2. Correlation analysis table

	y	x_1	x_2	x_3	x_4	x_5	x_6	x_7
y	1							
x_1	0.671^{***}	1						
x_2	0.713^{***}	0.644^{***}	1					
x_3	0.707^{***}	0.685^{***}	0.822^{***}	1				
x_4	0.739^{***}	0.747^{***}	0.758^{***}	0.973^{***}	1			
x_5	0.751^{***}	0.724^{***}	0.692^{***}	0.883^{***}	0.939^{***}	1		
x_6	0.716^{***}	0.585^{***}	0.651^{***}	0.809^{***}	0.814^{***}	0.841^{***}	1	
x_7	0.600^{***}	0.760^{***}	0.523^{***}	0.764^{***}	0.845^{***}	0.813^{***}	0.575^{***}	1

and apply the system of GMM estimation to eliminate endogenous for further analysis. In this paper, the regression model is as follows, model 1 is on behalf of the panel data model, and model 2 is on behalf of the dynamic panel model.

$$\text{Model 1} \quad y_{it} = \alpha_i + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_5 + \beta_4 x_6 + \mu_{it},$$
$$\text{Model 2} \quad y_{it} = \alpha_i + \gamma y_{it-1} + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_5 + \beta_4 x_6 + \mu_{it}.$$

Model 1 is the result of fixed effect, model 2 is he result of random effect. The result of Houseman test showed that Chi2(4) $=$ 7.22, the value of P is 0.1249, that can't refuse null hypothesis of Houseman test, so random effects model should be used. It can be seen from model 2, the technical innovation ability can really promote the upgrading of industrial structure for the manufacturing industry. In particular, sales revenue of new products increased by 1%, the industry's main business profit increased 0.222%; cost of technical renovation increased by 1%, the industry's main business profit increased 0.17%. Because of the diversification of products and competition, the invention of new products and the input of technological transformation can increase brand influence and competitiveness in a highly competitive market environment, thus increasing corporate profits. From the point of control variables, net value of fixed assets increased by 1%, the industry's main business profit increased 0.173%; average number of employees in manufacturing industry increased by 1%, the industry's main business profit increased 0.323%.

After the testing, the best model for dynamic panel is model 3 after the introduction of the second order. It is found that the equation is over identified and the value of P is 0.4130, which shows that the instrumental variable is effective, So the model is tenable. The lagged items of the explanatory variables entered into the model and the author found that the lagged items had a significant impact. Specifically, a lag of the main business profit increased by 1%, the current main business profit will be reduced by 0.342%; two lag of main business profit increased by 1%, the current main business profit will be reduced by 0.301%; it can be seen that three lag of main business profits will increase the current main business profit from the model 4. Perhaps it was related to the

Table 3. Regression result table

Variables	(1) fe y	(2) re y	(3) sysgmm5 y	(4) sysgmm6 y
L.y			-0.342^{***} (0.021)	-0.684^{***} (0.104)
L2.y			-0.301^{***} (0.0245)	-0.440^{***} (0.0631)
L3.y				1.862^{***} (0.482)
x1	0.100 (0.0857)	0.173^{***} (0.0643)	0.00873 (0.0667)	-1.061^{**} (0.539)
x2	−0.0677 (0.357)	0.323^{***} (−0.0981)	1.752^{***} (0.194)	2.011^{***} (0.531)
x5	0.503^{***} (0.151)	0.222^{***} (0.0813)	0.289^{***} (0.0488)	0.174 (0.151)
x6	0.148* (0.079)	0.170^{***} (0.0543)	0.447^{***} (0.0292)	−0.0328 (0.0508)
Constant	10.37^{**} (5.026)	9.734^{***} (1.694)	5.999^{***} (2.058)	1.966 (4.946)
Observations	270	270	166	114
R-squared	0.312			
Number of ind1	52	52	52	44

Standard errors in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$)

manufacturing industry in our country which is mainly processing enterprises, and many of which is low value-added industries. Companies tend to increase plant equipment investment in the current or in the next period to increase the profits in the future when the current operating profit increased (Table 3).

5 Conclusion

In this paper, it can be seen from the empirical results that the invention of new products and technological innovation can increase the main business profits and promote the upgrading of manufacturing industrial structure. China is a big manufacturing country, whose economic growth is mainly driven by the second industry at present, and China's manufacturing industry accounted for the largest proportion of the second industry. At the same time, the manufacturing industry can absorb a lot of labor force. The development of the manufacturing industry is of great significance to the increase of the employment rate, reduction of the gap between the rich and the poor and the stability of the society. Therefore, the government should support the technological innovation of manufacturing industry, the enterprise itself also should actively seek innovation to make themselves invincible in the tide of marketization and internationalization.

References

1. Adak M (2015) Technological progress, innovation and economic growth; the case of turkey. Procedia Soc Behav Sci 195:776–782
2. Castellacci F (2008) Innovation and the competitiveness of industries: comparing the mainstream and the evolutionary approaches. Technol Forecast Soc Change 75(7):984–1006
3. Cohen WM (2015) Innovation and technological change, economics of. Int Encycl Soc Behav Sci 2:160–168
4. Guan JC, Yam RCM et al (2006) A study of the relationship between competitiveness and technological innovation capability based on dea models. Eur J Oper Res 170(3):971–986
5. Hong J, Bea F (2016) Do government grants promote innovation efficiency in china's high-tech industries? Technovation 57:4–13
6. Sun S, Ye Q (2015) Mechanism and strategic choice of innovation driven manufacturing transformation. Ind Technol Forum 2:15–18
7. Wang ZH, Dong CT (2012) An analysis on the consistency between manufacturing structure and quality structure of labor force in china:and the discussion of shortage of migrant workers, skilled personnel, and difficulty of graduates employment. Popul Econ 16(3):363–368
8. Zhao X, Qiao M (2012) Research on the mode and path of shanghai promoting manufacturing upgrading with high technology. Econ Res Shanghai 2:63–69

A Crash Counts by Severity Based Hotspot Identification Method and Its Application on a Regional Map Based Analytical Platform

Xinxin Xu[1], Ziqiang Zeng[2,3(✉)], Yinhai Wang[3], and John Ash[3]

[1] School of Tourism and Economic Management, Chengdu University, Chengdu 610106, People's Republic of China
xuxinxin@cdu.edu.cn

[2] Uncertainty Decision-Making Laboratory, Sichuan University, Chengdu 610065, People's Republic of China
zengziqiang@scu.edu.cn

[3] Smart Transportation Applications and Research Laboratory, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA

Abstract. This paper aims to develop a crash counts by severity based hotspot identification method by extending the traditional empirical Bayes method to a generalized nonlinear model-based mixed multinomial logit approach. A new safety performance index and a new potential safety improvement index are developed by introducing the risk weight factor and compared with traditional indexes by employing four hotspot identification evaluating methods. The comparison results reveal that the new safety performance index derived by the generalized nonlinear model-based mixed multinomial logit approach is the most consistent and reliable method for identifying hotspots. Finally, a regional map based analytical platform is developed by expanding the safety performance module with the new safety performance index and potential safety improvement functions.

Keywords: Hotspot identification · Crash severity · Safety performance index · Potential safety improvement index · Regional map based analytical platform

1 Introduction

The identification of crash hotspots, also referred to as hazardous road locations, high-risk locations, accident-prone locations, black spots, sites with promise, or priority investigation locations, is the first step of the highway safety management process [10]. Hotspot identification (HSID) is of great importance to land transport authorities in their efforts to improve the safety of highways. There is a fairly extensive body of literature focused on methods for HSID [1,2,11].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_23

There are papers that discuss methods based on accident count or frequency [3], papers that employ both accident rate (AR) and rate quality control [13], and others that adopt the joint use of accident frequency and rate to flag sites with promise [7]. To correct for the regression-to-the-mean bias associated with typical HSID methods [4], some researchers have suggested using the empirical Bayes (EB) techniques [5]. This method combines clues from both the accident history of a specific site and expected safety of similar sites, and has the advantage of revealing underlying safety problems which otherwise would not be detected.

Some recent studies indicated that the severity of crashes should not be neglected in the hotspot identification (HSID) process [12]. The hotspots corresponding to high crash risk locations can be quite different when considering the crash frequency by different levels of crash severity. It is particularly important to take into account crash severities in site ranking, because the cost of crashes could be hugely different at different severity levels. This means that, for instance, a road segment with higher frequency of fatal accidents may be considered more hazardous than a road segment with fewer fatal accidents, but more severe or minor injury accidents. Therefore, it is necessary to consider crash severity when identifying hotspots.

While the equivalent property damage only (EPDO) method [15] is a way of comparing severity types among each other, inconsistencies can occur when evaluating the HSID methods in different time periods, since the traditional EPDO method overemphasizes sites with a low frequency of fatal or severe crashes [10]. As a result, a risk weight factor is developed in this research by combining the average crash cost with the corresponding probability for each type of crash severity. In order to develop a new safety performance index (SPI) and a new potential safety improvement index (PSII), a generalized nonlinear model-based mixed multinomial logit approach is introduced to extend the traditional empirical Bayes method for estimating the probability of crashes in different severity levels. The new method developed in this paper is compared with other hotspot identification methods by employing four hotspot identification evaluating methods and applied on a regional map based analytical platform.

2 Crash Counts by Severity Prediction Model

The EB method requires the use of pertinent crash prediction models. For the purpose of the study, in order to account for unobserved heterogeneity and nonlinear effects of variables to extract more complex relationship, a refinement of the generalized nonlinear model-based (GNM-based) mixed multinomial logit (MNL) approach developed by Zeng et al. [17] was used.

2.1 GNM-based Mixed MNL Model

In this research, three categories are considered for the crash severity (i.e., property damage only (PDO) ($k = 1$), injury ($k = 2$), and fatal ($k = 3$)). The crash severity type, denoted by Y, was the response variable, whereas contributing factors for roadway geometric characteristics, traffic characteristics, and weather conditions were the independent variables denoted by x_{ij} $(j = 1, 2, \cdots, J)$, where i denotes the observation and J denotes the number of independent variables. Y is then defined as follows:

$$Y = \begin{cases} 1, & \text{if crash is PDO,} \\ 2, & \text{if crash is injury,} \\ 3, & \text{if crash is fatal,} \end{cases} \tag{1}$$

Ye and Lord [16] demonstrated that fatal crashes should be set as the baseline severity for the mixed MNL model. To minimize the bias and reduce the variability of a model, in this paper, fatal crashes were used as the baseline severity category for comparison with the other categories.

In order to account for the unobserved heterogeneity, let $\Omega = (\omega_1, \omega_2, \omega_3)$, as discussed previously, and note that the Ω vector has a continuous density function $f(\Omega|\Gamma)$, where Γ is a vector of parameters charactering the density function, $\omega_k = [\omega_{k1}, \omega_{k2}, \cdots, \omega_{kJ}]^T$ is the coefficient vector for the kth category of the predictor vector. According to [6,9,14], the resulting mixed MNL crash severity probabilities are as follows:

$$Pr(Y_i = K) = \int \frac{1}{1 + \sum_{k=1}^{K-1} e^{U_{ki}\omega_k + \beta_{k0}}} f(\Omega|\Gamma) d\Omega, i = 1, 2, \cdots, n, \tag{2}$$

$$Pr(Y_i = k) = \int \frac{e^{U_{ki}\omega_k + \beta_{k0}}}{1 + \sum_{k=1}^{K-1} e^{U_{ki}\omega_k + \beta_{k0}}} f(\Omega|\Gamma) d\Omega, i = 1, 2, \cdots, n; k = 1.2, \cdots, K-1, \tag{3}$$

where $Pr(Y_i = k)$ is the probability of crash severity type k; $U_{ki} = [U_{ki1}(x_{i1}), U_{ki2}(x_{i2}), \cdots, U_{kiJ}(x_{iJ})]$ is the nonlinear predictor vector of observation i for contributing factors (i.e., roadway geometric characteristics, traffic characteristics, weather conditions); β_{k0} is an intercept term specific to crash severity type k. Since U_{ki} is considered as a nonlinear predictor vector of observation i for contributing factors, Eqs. (2) and (3) are called the prediction functions of the GNM-based mixed MNL approach.

2.2 Predicting Crash Counts in Different Severity Levels

Based on the statistical approaches available to address the unobserved heterogeneity [8], the expected crash density in a roadway segment i can be estimated as below:

$$d_i = \int e^{U_i \omega + \beta_0} f(\omega|\varphi) d\omega, i = 1, 2, \cdots, n, \tag{4}$$

where $d_i = \mu_i / L_i y_i$ is the crash density of roadway segment i during a certain time period; μ_i is the expected crash frequency (or rate) along segment i during a certain time period; L_i is the segment length in miles; y_i is the time period length (years) of crash frequency of roadway segment i; $\omega = [\omega(1), \omega(2), \cdots, \omega(J)]^T$ is the coefficient vector for $U_i = [U_1(x_{i1}), U_2(x_{i2}), \cdots, U_J(x_{iJ})]$ when estimating the expected crash density.

According to Eqs. (2)–(4), the expected crash density for different severity levels can be estimated as follows.
(1) Expected PDO crash density:

$$\begin{aligned} d_{i1} &= d_i \cdot Pr(Y_i = 1) \\ &= \int e^{U_i\omega+\beta_0} f(\omega|\varphi)d\omega \cdot \int \frac{e^{U_{1i}\omega_1+\beta_{10}}}{1+\sum_{k=1}^{2} e^{U_{ki}\omega_k+\beta_{k0}}} f(\Omega|\Gamma)d\Omega,\ i = 1, 2, \cdots, n, \end{aligned} \tag{5}$$

where d_{i1} is the expected PDO crash density along segment i during a certain time period.
(2) Expected injury crash density:

$$\begin{aligned} d_{i2} &= d_i \cdot Pr(Y_i = 2) \\ &= \int e^{U_i\omega+\beta_0} f(\omega|\varphi)d\omega \cdot \int \frac{e^{U_{2i}\omega_2+\beta_{20}}}{1+\sum_{k=1}^{2} e^{U_{ki}\omega_k+\beta_{k0}}} f(\Omega|\Gamma)d\Omega,\ i = 1, 2, \cdots, n, \end{aligned} \tag{6}$$

where d_{i2} is the expected injury crash density along segment i during a certain time period.
(3) Expected fatal crash density:

$$\begin{aligned} d_{i3} &= d_i \cdot Pr(Y_i = 3) \\ &= \int e^{U_i\omega+\beta_0} f(\omega|\varphi)d\omega \cdot \int \frac{1}{1+\sum_{k=1}^{2} e^{U_{ki}\omega_k+\beta_{k0}}} f(\Omega|\Gamma)d\Omega,\ i = 1, 2, \cdots, n, \end{aligned} \tag{7}$$

where d_{i3} is the expected fatal crash density along segment i during a certain time period. Based on Eqs. (5)–(7), a crash counts by severity based hotspot identification (CCS-based HSID) method is developed in the following section.

3 Crash Counts by Severity Based Hotspot Identification Method

In this section, a CCS-based HSID method is developed by employing the GNM-based MNL approach discussed above. A new SPI and a new PSII are developed by introducing the risk weight factor.

3.1 Safety Performance Index

Based on Eqs. (5)–(7), the EPDO crash frequency measure is modified and employed to weight crashes according to severity (fatal, injury, and PDO) to develop a combined crash density and severity score (CCDSS) for each site [15]. The weight factors are based on PDO crash costs. An EPDO value summarizes the crash costs and severity.

In the calculations, weight factors were assessed from the crash cost estimates developed by WSDOT in the Annual Collision Data Summary Reports (2011–2014). Using average crash costs for motorways, fatal crashes ($2,227,851) have a weight factor equal to 981, injury crashes ($20,439) have a weight factor equal to 9, and PDO crashes ($2,271) have a weight factor equal to 1. However, if we only consider the average crash costs to be the weight factor, inconsistencies can occur when evaluating the HSID methods in different time periods, since the traditional EPDO method overemphasizes sites with a low frequency of fatal or severe crashes [10]. As a result, a risk weight factor is developed in this research by combining the average crash cost with the corresponding probability for each type of crash severity. Let F_w denote the fatality risk weight factor, I_w, the injury risk weight factor, and P_w, the PDO risk weight factor, then they are defined by using the following equations:

$$F_w = \frac{c_F \cdot \eta_F}{c_P \eta_P}, \;\; I_w = \frac{c_I \cdot \eta_I}{c_P \eta_P}, \;\; P_w = 1, \tag{8}$$

where $c_F = \$2,227,851$, $c_I = \$20,439$, and $c_p = \$2,271$ are the average costs for fatal, injury, and PDO crashes; η_F, η_I, and η_P are the probabilities of occurrence for fatal, injury, and PDO crashes.

Based on the preceding analysis, the expected CCDSS (ECCDSS) for roadway segment i can be defined as:

$$\text{ECCDSS}_i = d_{i1} F_w + d_{i2} I_w + d_{i3} P_w, \;\; i = 1, 2, \cdots, n. \tag{9}$$

Equation (9) is regarded as a safety performance function (SPF). In fact, the ECCDSS is an extension of the expected crash density based on the GNM. When the fatality weight factor and the injury weight factor are both equal to 1 (i.e., $F_w = I_w = 1$), the ECCDSS will degrade to an expected crash density based on the GNM. The EB method is a statistical method that combines the observed crash frequency with the predicted crash frequency using the SPF to calculate the expected crash frequency for a site of interest. The EB method pulls the crash count towards the mean, accounting for the regression to the mean (RTM) bias. A lot of studies have proved that the EB approach is the most consistent and reliable method for identifying sites with promise [2,10]. In this research, the EB method is employed to develop the new SPI as shown in the following:

$$\text{SPI}_i = \lambda_i \text{ECCDSS}_i + (1 - \lambda_i) \text{OCCDSS}_i, \;\; i = 1, 2, \cdots, n, \tag{10}$$

where OCCDSS_i is the observed combined crash density and severity score (OCCDSS) for roadway segment i and is defined as below:

$$\text{OCCDSS}_i = \sigma_{i1} F_w + \sigma_{i2} I_w + \sigma_{i3} P_w, \;\; i = 1, 2, \cdots, n, \tag{11}$$

where σ_{i1}, σ_{i2}, σ_{i3} are the observed fatal, injury, and PDO crash density along segment i during a certain time period respectively; λ_i is a weighting factor that is calculated through the following equation:

$$\lambda_i = \frac{1}{1 + \alpha_i \text{ECCDSS}_i}, \tag{12}$$

where α_i is the over dispersion parameter, which is a constant for a given model and is derived during the regression calibration process.

3.2 Potential Safety Improvement Index

The PSII was developed as the difference between the SPI and the ECCDSS, as follows:

$$\begin{aligned}\mathrm{PSII}_i &= \lambda_i \mathrm{ECCDSS}_i + (1-\lambda_i)\mathrm{OCCDSS}_i - \mathrm{ECCDSS}_i \\ &= \mathrm{SPI}_i - \mathrm{ECCDSS}_i,\ \ i = 1, 2, \cdots, n,\end{aligned} \tag{13}$$

when the PSII value is greater than zero, a site experiences a higher combined frequency and severity score than expected; when the PSII value is less than zero, a site experiences a lower combined frequency and severity score than expected.

4 Case Study Data

This study was performed based on crash data records collected in Washington State from January 2011 to December 2014 (i.e., a four-year period). The data were obtained from the Washington State Department of Transportation (WSDOT), Highway Safety Information System (HSIS), and the Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net) platform at the University of Washington (UW). Four major datasets are included in this study: crash data, roadway geometric characteristics, traffic characteristics, and weather conditions. These datasets detail all of the information regarding crash frequency, locations, severities, roadway segment length, average number of lanes (NOL), horizontal curve type (HCT), curvature of the segment (COS), average width of outer shoulder (WOS), average width of inner shoulder (WIS), average width of median (WM), dominant lane surface type (DLST), dominant outer shoulder type (DOST), dominant inner shoulder type (DIST), dominant median type (DMT), average speed limit (ASL), AADT, AADT per lane, road surface conditions (RSC, i.e., dry, wet, snow/ice/slush), and visibility (good, bad).

In this research, we consider using the proportion of the crash frequency for each type of severity based on the collected crash data of 21,396 road segments along I-5, I-90, I-82, I-182, I-205, I-405 and I-705 in Washington to represent the probability of crash occurrence for each severity level in this area. Based on the crash counts after data quality control, the total number of crashes recorded during the data collection period was 47,657, including 134 fatal crashes, 13,824 injury crashes, and 33,699 PDO crashes. Thus, we can calculate that $\eta_F = 0.0028$, $\eta_I = 0.29$, $\eta_P = 0.7072$; then, the values of the risk weight factors are obtained by employing Eq. (8) as $F_w = 3.884$, $I_w = 3.691$, $P_w = 1$.

5 Test Results and Discussion

In order to demonstrate the effectiveness of the SPI and PSII as developed in this research, they are compared with six reference performance indexes, which include the expected crash density based on the conventional safety performance function from the Highway Safety Manual (HSM) (i.e., NB GLM), expected crash density based on the GNM, EB estimated crash density based on the NB GLM,

Table 1. Results of site consistency test of various HSID methods

HSID method index (h)	HSID method name	Parameter Settings			
		$\gamma = 0.01$		$\gamma = 0.05$	
		$SCT_{h,t}$ Period 1 (2011–2012)	$SCT_{h,t+1}$ Period 2 (2013–2014)	$SCT_{h,t}$ Period 1 (2011–2012)	$SCT_{h,t+1}$ Period 2 (2013–2014)
1	SPI	22943.06	21521.79	46310.12	44251.68
2	PSII	22198.11	20879.23	45389.79	43921.56
3	CD (NB GLM)	21036.33	20105.38	43897.54	42868.83
4	CD (GNM)	21901.86	20993.89	44890.67	43571.45
5	EB CD (NB GLM)	21896.77	20981.31	45303.46	43784.09
6	EB CD (GNM)	22856.34	21349.04	45987.98	44012.49
7	ARP (NB GLM)	20131.67	19034.25	43015.82	42106.73
8	ARP (GNM)	20358.78	19734.55	43823.17	42871.29

EB estimated crash density based on the GNM, ARP based on the NB GLM, and ARP based on the GNM.

Cheng and Washington [2] have developed four new evaluation tests for HSID. In this research, the site consistency test, method consistency test, total rank differences test, and the total score test are employed to evaluate the effectiveness of the developed safety performance indexes and reference performance indexes.

The evaluation experiment uses the following procedure, which closely mimics how reactive safety management programs are conducted in practice:

(1) For the purpose of comparing alternate HSID approaches, the 4-year accident data were separated into two periods, Period 1 (Year 2011–2012) and Period 2 (Years 2013–2014).
(2) Road sections (intersections, ramps, two-lane rural roads, etc.) are segregated so that the safety of similar sites can be fairly compared. In this evaluation, the analysis is based on the analysis of nine functional classifications of road sections.
(3) For each HSID method, similar road sections are sorted in descending order of estimated safety (noting that the four HSID methods rank sites according to different criteria).
(4) Sections with the highest rankings are flagged as hotspots (in practice these sites will be further scrutinized). Typically, a threshold is assigned according to safety funds available for improvement, such as the top 1% of sites. In this evaluation, both the top 1% and 5% of the locations are used as experimental values.

5.1 Site Consistency Test

The site consistency test (SCT) measures the ability of an HSID method to consistently identify a high-risk site over repeated observation periods. The test rests on the premise that a site identified as high risk during time period i should also reveal an inferior safety performance in a subsequent time period $t+1$, given that the site is in fact high risk and no significant changes have occurred at the site. The method that identifies sites in a future period with the highest crash frequency is the most consistent. In this research, the SPI developed above is employed as the safety performance criterion in the subsequent time period. The test statistic is given as:

$$\mathrm{SCT}_{h,t+1} = \sum_{q=n-n\gamma+1}^{n} \mathrm{SPI}_{q,h,t+1},\ h = 1, 2, \cdots, H, \tag{14}$$

where h is the HSID method index being compared; n is the total number of roadway segment, γ is the threshold of identified hotspots (e.g., $\gamma = 0.01$ corresponds with top 1% of n roadway segments identified as hotspots, and $n\gamma$ is the number of identified hotspots).

From the site consistency test, it is shown in Table 1 that the SPI method outperforms other HSID methods in identifying both of the top 1% and 5% of hotspots with highest SCT values, 21521.79 and 44251.68, in Period 2, followed closely by the EB CD (GNM) method. The ARP (NB GLM) performs the worst in both cases, with the identified hotspots experiencing the lowest number of SCT values, say, 19034.25 and 42106.73, respectively (although the ARP is based on reduction potential, so the total count can be misleading).

5.2 Method Consistency Test

The method consistency test (MCT) evaluates a method's performance by measuring the number of the same hotspots identified in both time periods. It is assumed that road sections are in the same or similar underlying operational state and their expected safety performance remains virtually unaltered over the two analysis periods. With this assumption of homogeneity, the greater the number of hotspots identified in both periods the more consistent the performance of the HSID method. The test statistic is given as:

$$\mathrm{MCT}_h = \{s_{n-n\gamma+1}, s_{n-n\gamma}, \cdots, s_n\}_{h,t} \bigcap \{s_{n-n\gamma+1}, s_{n-n\gamma}, \cdots, s_n\}_{h,t+1},\ h = 1, 2, \cdots, H, \tag{15}$$

here, only segments $\{s_{n-n\gamma+1}, s_{n-n\gamma}, \cdots, s_n\}$ identified in the top threshold γ are compared. Table 2 shows the number of similarly identified hotspots identified by alternate HSID methods over the two periods. The SPI method is superior in this test by identifying the largest number of the same hotspots in both cases of $\gamma = 0.01$ and $\gamma = 0.05$, with 124 and 546 roadway segments, respectively. In other words, the SPI method identified 124 segments in 2011–2012 that were also identified as hotspots in 2013–2014. The ED CD (GNM), which performs slightly better than the ED CD (GLM) method, places 2nd with identifying 118 consistent hotspots (in the case of $\gamma = 0.01$) and 472 consistent hotspots (in the case of $\gamma = 0.01$). The ARP (NB GLM) performs worst, with the lowest number of consistent hotspots identified in the two periods. Again, the SPI method outperforms the other HSID methods. Also shown in Table 2 are differences between percentages (shown in the parentheses) of Column 3 and Column 4 for the eight methods. There is a consistent drop in percentages as threshold values drop. The explanation is that the top segments suffer from greater random fluctuations in crashes, and thus the higher is the threshold, the larger are the random fluctuations and the likelihood of not being identified in a prior period.

Table 2. Results of method consistency test of various HSID methods

HSID method index (h)	HSID method name	Parameter settings	
		$\gamma = 0.01$	$\gamma = 0.05$
1	SPI	124 (60.7%)	546 (53.5%)
2	PSII	103 (50.5%)	452 (44.3%)
3	CD (NB GLM)	92 (45.1%)	406 (39.8%)
4	CD (GNM)	101 (49.5%)	423 (41.5%)
5	EB CD (NB GLM)	109 (53.4%)	441 (43.2%)
6	EB CD (GNM)	118 (57.8%)	472 (46.3%)
7	ARP (NB GLM)	83 (40.7%)	382 (37.5%)
8	ARP (GNM)	88 (43.1%)	393 (38.5%)

5.3 Total Rank Differences Test

The total rank differences test (TRDT) takes into account the safety performance rankings of the road sections in the two periods. The test is conducted by calculating the sum of the total rank differences of the hotspots identified across the two periods. The smaller the total rank difference, the more consistent the HSID method. The test statistic is given as:

$$\mathrm{TRDT}_h = \sum_{q=n-n\gamma+1}^{n} |R(q_{h,t}) - R(q_{h,t+1})|, \;\; h = 1, 2, \cdots, H, \tag{16}$$

where $R(q_{h,t})$ is the rank of segment q in period t for method h. The difference in ranks is summed over all identified segments for threshold level γ for period t. Table 3 illustrates that the SPI method is superior in the total rank differences test. In both the $\gamma = 0.01$ and $\gamma = 0.05$ cases, the SPI method has significantly smaller-summed ranked differences, by about 22.6% (in the case of $\gamma = 0.01$) and 16.9% (in the case of $\gamma = 0.05$) compared with the EB CD (GNM), and by about 75.1% (in the case of $\gamma = 0.01$) and 77.4% (in the case of $\gamma = 0.05$) compared with the ARP (NB GLM). This result suggests that the SPI method is the best HSID method (of the eight evaluated here) for ranking roadway segments consistently from period to period.

5.4 Total Score Test

The total score test (TST) combines the site consistency test, the method consistency test, and the total rank difference test in order to provide a synthetic index. The test statistic is given as:

$$\mathrm{TST}_h = \frac{100}{3}\left[\left(\frac{\mathrm{SCT}_{h,t+1}}{\max_h\{\mathrm{SCT}_{h,t+1}\}}\right) + \left(\frac{\mathrm{MCT}_h}{\max_h\{\mathrm{MCT}_h\}}\right) + \left(1 - \frac{\mathrm{TRDT}_h - \min_h\{\mathrm{TRDT}_h\}}{\max_h\{\mathrm{TRDT}_h\}}\right)\right],$$
$$h = 1, 2, \cdots, H, \tag{17}$$

where the test assumes that the SCT, MCT, and TRDT have the same weight. The former three tests provide absolute measures of effectiveness, whereas the total score test gives an effectiveness measure relative to the methods being compared. If method h performed best in all of the previous tests, the TST value is equal to 100. If method h performed worst in all of the tests, the TST value is positive since all three components of the test have a positive value. Indeed, SCT and MCT, which should be maximized by the HSID methods, are weighted

Table 3. Results of total rank differences test of various HSID methods

HSID method index (h)	HSID method name	Parameter settings	
		$\gamma = 0.01$	$\gamma = 0.05$
1	SPI	2354	10237
2	PSII	3031	12798
3	CD (NB GLM)	3672	14781
4	CD (GNM)	3298	13587
5	EB CD (NB GLM)	3158	13016
6	EB CD (GNM)	2887	11973
7	ARP (NB GLM)	4123	18167
8	ARP (GNM)	3887	17105

Table 4. Results of total score test of various HSID methods

HSID method index (h)	HSID method name	Parameter settings	
		$\gamma = 0.01$	$\gamma = 0.05$
1	SPI	100	100
2	PSII	87.89	89.31
3	CD (NB GLM)	78.55	82.07
4	CD (GNM)	85.37	85.83
5	EB CD (NB GLM)	88.63	88.14
6	EB CD (GNM)	93.81	92.12
7	ARP (NB GLM)	70.82	73.82
8	ARP (GNM)	75.16	77.02

in relation to the maximum values in the tests, whereas TRDT, which should be minimized by the HSID methods, is weighted in relation to its difference from the minimum value in the test. Table 4 illustrates the results of total score test of the eight HSID methods, in which SPI performed best in both $\gamma = 0.01$ and $\gamma = 0.05$ cases, and was followed closely by EB CD (GNM) method with 93.81 score (in the case of $\gamma = 0.01$) and 92.12 score (in the case of $\gamma = 0.05$). ARP (NB GLM) performed the worst in both cases, with 70.82 score and 73.82 score respectively.

Overall, the four tests reveal that the SPI method is the most consistent and reliable method for identifying hotspots. Although it can only be applied to roadway segments where the crash data for different levels of severity are available, with the rapid development of intelligent transportation systems and data collection technologies, this method could become quite useful in identifying high-risk road sites. On several criteria, the SPI outperforms other methods by a wide margin. This evaluation suggests that the SPI method (of the methods compared) has a potential to become the industry standard.

6 Application on Regional Map Based Analytical Platform

A regional map based analytical platform was developed on the DRIVE Net system to highlight the methodology developed under this project. Ultimately, the existing safety performance analysis function under the "Safety Performance" module was expanded. The SPI developed in the preceding is used to color-code the regional map based on safety performance. The PSII is employed to highlight potential safety improvements on the map. By combining the two indices on the regional map, one can easily identify accident hotspots and the key influencing factors to consider in an improvement package.

The interface of the safety performance module in the regional map based analytical platform is illustrated in Fig. 1. There are three sub-functions implemented on this panel: Incident Frequency (NB GLM), Estimated Crash Mean and Potential Safety Improvement Index (ARP NB GLM). The new SPI and PSII were added as expended safety performance analysis options. As stated in the previous modeling part of this report, within a selected time range and corridor, the SPI shows a more comprehensive view of safety performance on a given corridor. The accident/incident data is from Washington Incident Tracking System (WITS) database. The SPI level ranges from Level A to Level F, where Level A (light green) corresponds to the highest safety performance and Level F (dark red) corresponds to the lowest safety performance expected as shown in Fig. 1.

The PSII implements the EB method in the modeling part. In this function, both the historical incident data and the characteristics of the selected corridor are used as model inputs. The output format still uses the six different colors representing Level A to Level F to show the potential safety improvement index on the map, where Level A shows the segment has the least potential to improve

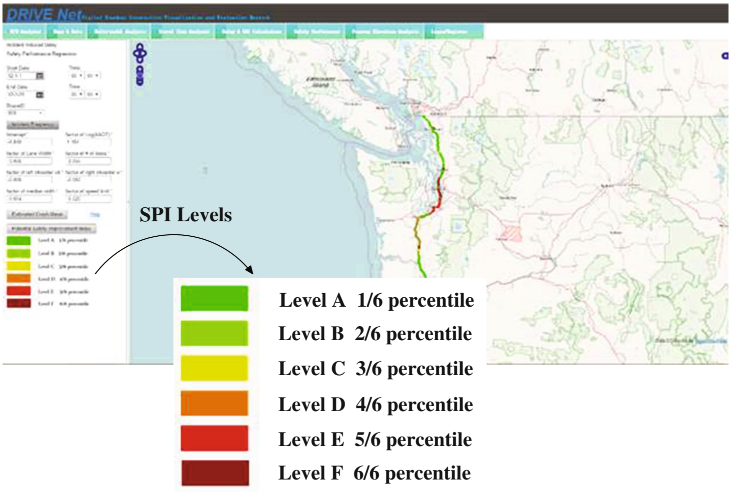


Fig. 1. SPI level ranges from Level A to Level F in the safety performance module

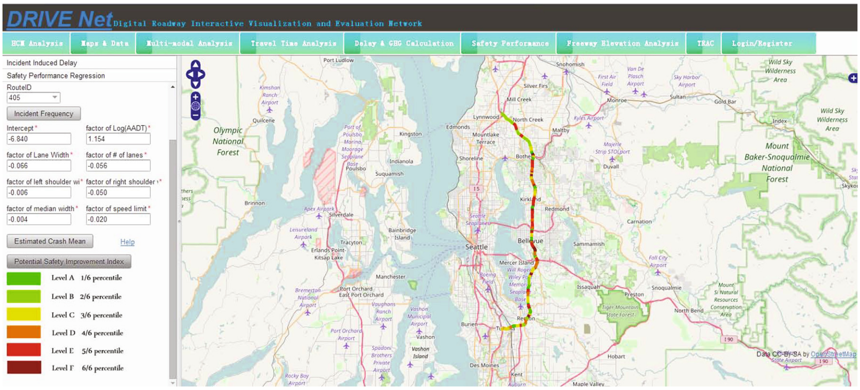


Fig. 2. An example of the PSII function

its safety, and Level F shows the segment has the most potential to improve its safety. Figure 2 shows an example of this function.

7 Conclusions

A CCS-based HSID method is developed by extending the traditional EB method to a GNM-based mixed MNL approach in this paper. A new SPI and a new PSII are developed by introducing the risk weight factor and compared with traditional indexes by employing four HSID evaluating methods, including the site consistency test, method consistency test, total rank differences test, and the total score test. The test results showed that the new SPI derived by the

GNM-based mixed MNL approach is the most consistent and reliable method for identifying hotspots. Finally, the new CCS-based HSID method was applied on a regional map based analytical platform.

Acknowledgements. This research was supported by the Youth Program of National Natural Science Foundation of China (Grant No. 71501137), the General Programs of China Postdoctoral Science Foundation (Grant No. 2015M572480), the International Postdoctoral Exchange Fellowship Program of China Postdoctoral Council (Grant No. 20150028), the project of Research Center for System Sciences and Enterprise Development (Grant No. Xq16B05), and also funded by Sichuan University (Grant No. skqy201647). The authors would like to give our great appreciation to the editors and anonymous referees for their helpful and constructive comments and suggestions, which have helped to improve this article.

References

1. Cheng W, Washington S (2005) Experimental evaluation of hotspot identification methods. Accid Anal Prev 37:870–881
2. Cheng W, Washington S (2008) New criteria for evaluating methods of identifying hot spots. Transp Res Rec 2083:76–85
3. Deacon J, Zegeer C, Deen R (1975) Identification of hazardous rural highway locations. Transp Res Rec 543:16–33
4. Hauer E (1980) Bias-by-selection: overestimation of the effectiveness of safety countermeasures caused by the process of selection for treatment. Accid Anal Prev 12:113–117
5. Hauer E, Persaud B, Smiley A et al (1991) Estimating the accident potential of an Ontario driver. Accid Anal Prev 23:133–152
6. Hensher D, Greene W (2003) The mixed logit model: the state of practice. Transportation 30:133–176
7. Laughland J, Haefner L, Hall J et al (1975) Methods for evaluating highway safety improvements. NCHRP 162, Transportation Research Board
8. Mannering F, Shankar V, Bhat C (2016) Unobserved heterogeneity and the statistical analysis of highway accident data. Anal Methods Accid Res 11:1–16
9. McFadden D, Train K (2000) Mixed MNL models for discrete response. J Appl Econometrics 15:447–470
10. Montella A (2010) A comparative analysis of hotspot identification methods. Accid Anal Prev 42:571–581
11. Park B, Lord D, Lee C (2014) Finite mixture modeling for vehicle crash data with application to hotspot identification. Accid Anal Prev 71:319–326
12. Qu X, Meng Q (2014) A note on hotspot identification for urban expressways. Saf Sci 66:87–91
13. Stokes R, Mutabazi M (1996) Rate-quality control method of identifying hazardous road locations. Transp Res Rec 1542:44–48
14. Train K (2003) Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge
15. Washington S, Haque M, Oh J et al (2014) Applying quantile regression for modeling equivalent property damage only crashes to identify accident blackspots. Accid Anal Prev 66:136–146

16. Ye F, Lord D (2011) Investigation of effects of underreporting crash data on three commonly used traffic crash severity models multinomial logit, ordered probit, and mixed logit. Transp Res Rec 2241:51–58
17. Zeng Z, Zhu W, Ke R et al (2017) A generalized nonlinear model-based mixed multinomial logit approach for crash data analysis. Accid Anal Prev 99:51–65

Comparison Between K-Means and Fuzzy C-Means Clustering in Network Traffic Activities

Purnawansyah[1], Haviluddin[2(✉)], Achmad Fanany Onnilita Gafar[3], and Imam Tahyudin[4]

[1] Faculty of Computer Science, Universitas Muslim Indonesia, Makassar, Indonesia
[2] Faculty of Computer Science and Information Technology, Mulawarman University, Samarinda, Indonesia
haviluddin@gmail.com
[3] State Polytechnic of Samarinda, Samarinda, Indonesia
[4] Information System Department, STMIK AMIKOM, Purwokerto, Indonesia

Abstract. A network traffic utilization in order to support teaching and learning activities are an essential part. Therefore, the network traffic management usage is requirements. In this study, analysis and clustering network traffic usage by using K-Means and Fuzzy C-Means (FCM) methods have been implemented. Then, both of method were used Euclidean Distance (ED) in order to get better results clusters. The results showed that the FCM method has been able to perform clustering in network traffic.

Keywords: Network traffic · K-Means · Fuzzy C-Means · Clustering

1 Introduction

A bandwidth usage monitoring management in a network traffic at university is indispensable. Therefore, a recording and analysis of bandwidth usage by network administrators is very beneficial. It aims to make use of the bandwidth can be well controlled, stable access, and gives users convenient access. Therefore, mapping or cluster of bandwidth usage in order to support the network administrator performance analysis is required.

In this study, the cluster method based on intelligence algorithm are proposed in bandwidth usage, such as Self-Organizing Maps (SOM), K-Means, Fuzzy C-Means, etc. For that reason, these algorithms have been widely used by researchers in order to solve cluster data problems in a variety of fields, including economics [5], supply chain [2], engineering [6], hydrology [1,4], internet and social media [3,8], pattern recognition [7], and so forth. Many research results have shown that the algorithms are able to provide accurate information in solving the clustering problem.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_24

Afterward, two intelligence algorithms is K-Means and Fuzzy C-Means (FCM) for cluster bandwidth usage data has been implemented. The purpose of this study are compared the feasibility of two clustering methods and how it works in the real world problems particularly subject on network traffic. Thus, the expected results of network traffic clustering are coordination schemes that support resource management [8]. Furthermore, this paper will apply two models, namely K-Means and Fuzzy C-Means (FCM) that have been developed and compared in order to cluster the network traffic usage. Section 2 describes the architectures of K-Means and FCM clustering models are proposed. Section 3 describes the analysis and discussion of the results. Finally, conclusions are summarized in Sect. 4.

2 Research Method

In this section, a brief information of K-Means and FCM models are presented.

2.1 Principle of K-Means Method

K-means clustering is an unsupervised learning classified. This algorithm is based on the determination of the distance between the centroid and the training data. Then, the number of cluster centroid based on the number desired. Meanwhile, the initialization centroid randomly generated by considering the data training. In other words, the centroid should be in the training data space. Then, a couple of training data is from the attributes of the data patterns to be planned. In each iteration, the distance of each training data with each centroid to be calculated. It means that any training data will have a centroid distance. At the same time, members of the cluster indicated by the smallest distance from the corresponding centroid. Then, a new centroid value is calculated based on the average value of each member of the cluster. If the cluster member does not change then the iteration is stopped, Fig. 1.

The following data clustering techniques using the K-Means algorithm as follows.

- Determine the number of clusters K.
- The Initialization of K cluster centers can be done randomly and used as initial cluster centers.
- Allocate all data/objects to the nearest cluster. The proximity of the two objects is determined by the distance of the object. To calculate the distance of all the data to the cluster center point using the Euclidean distance theory formulated as follows:

$$T_{(x,y)} = \sqrt{(T_{1x} - T_{1y})^2 + (T_{2x} - T_{2y})^2 + \cdots + (T_{kx} - T_{ky})^2}, \tag{1}$$

where, $T_{(x,y)}$ is distance data of x to the cluster center of y; T_{kx} is i-data on attribute data of k; T_{ky} is the center point of j on the attribute of k.

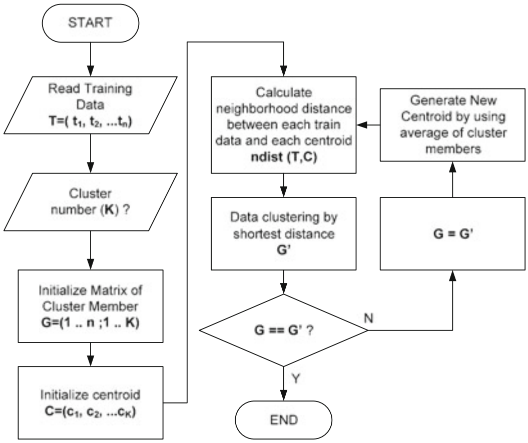


Fig. 1. K-Means algorithm

- Recalculate the center of the cluster with the new cluster membership. This is computed by determining the centroid/center cluster.
- Assign each object to put on the new cluster center, if the center of the cluster changed, back to 3, otherwise clustering is complete.
- Analyze the results in the clustering process.

2.2 Principle of Fuzzy C-Means Method

Other advanced techniques clustering in machine learning model come from the Dunn in 1973, and improved by Bezdek in 1981, called Fuzzy C-Means (FCM) clustering was developed by Dunn in 1973, and improved by Bezdek in 1981. In principle, FCM clustering process is based on a partition of a set of data into a number of clusters with minimum similarity between different clusters [2]. Since the introduction of the fuzzy set theory in 1965 by Zadeh, it has been applied in a variety of fields. FCM is a flexible fuzzy partition that an improvement of common C-Means algorithm [5]. At FCM, each feature vectors valued between [0–1] by using the membership function, because FCM is based on the criteria of distances numbers between clusters. In other words, FCM clustering is based on Point-Prototype Clustering Model with output centroid most optimal partition. Where, partition optimization centroid is obtained by minimizing the objective function. The formula is given by:

$$J_{\mathrm{FCM}}(U,V) = \sum_{j=1}^{N}\sum_{i=1}^{C}(u_{ij})^{q}(d_{ji})^{2},$$

where
$U =$ Fuzzy datasets K-partition;
$V =$ Set of prototype centroid;
$V = \{v_1, v_2, \cdots, v_C\} \subset R^P$;

$$(d_{ji})^2 = \|x_j - v_i\|^2 = \sqrt{\left(x_{j\,(\text{row})} - v_{i\,(\text{row})}\right)^2 + \left(x_{j\,(\text{col})} - v_{i\,(\text{col})}\right)^2}. \tag{2}$$

Euclidean Distance between x_j and v_i;
$X = \{x_1, x_2, \cdots, x_n\} \subset R^P$;
$v_i =$ Centroid cluster to i;
$u_{ij} =$ Membership level x_j in cluster to i;
$N =$ Total data;
$C =$ Total cluster;
$q =$ Fuzzifier parameter, $q > 1$.

Afterward, all of the objects in each cluster has a certain degree of proximity or similarity. Meanwhile, the FCM processes consists of five stages. First, to determine the cluster is set become a center cluster location marker on average for each cluster. Second, calculate the distance between feature vector (X) and the centroid vector $(V)[X \to V]$. In this experiment, Euclidean Distance (ED) was implemented. Third, Calculate membership level. Fourth, Calculate new centroid. Lastly, recalculated the new centroid, if criterion between [0–1] is reached then stop the iteration. In this study, the FCM Clustering algorithm is as follows:

Step 1. Initialization Vector centroid, v_i (prototypes).
Step 2. Calculate the distance between feature vector (X) and the centroid vector $(V)[X \to V]$. Feature vectors with the closest distance to one of the centroid vectors then expressed as a cluster member.
Step 3. Calculate membership level of all feature vectors in all clusters by using the formula:

$$u_{ij} = \frac{1}{\sum_{k=1}^{K}\left[\frac{(d_{ji})^2}{(d_{jk})^2}\right]^{1/(q-1)}} = \frac{\left[\frac{1}{(d_{ji})^2}\right]^{1/(q-1)}}{\sum_{k=1}^{K}\left[\frac{1}{(d_{jk})^2}\right]^{1/(q-1)}}. \tag{3}$$

Step 4. Calculate new centroid using Eq. (4).

$$\hat{V_i} = \frac{\sum_{j=1}^{N}(u_{ij})^q X_j}{\sum_{j=1}^{N}(u_{ij})^q}. \tag{4}$$

Step 5. Recalculate the step 4, $u_{ij} \to \hat{u}_{ij}$ If, $\max_{ij}|u_{ij} - \hat{u}_{ij}| < \varepsilon$, where ε is termination criteria between 0 and 1. Then, the iteration process is stopped. If not go back to step 5. The FCM algorithm can be seen in Fig. 2.

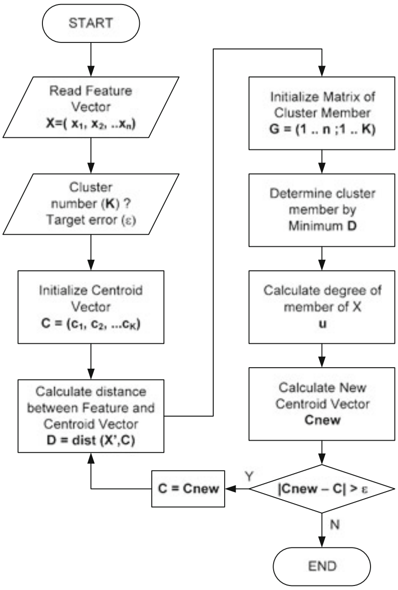


Fig. 2. Fuzzy C-means algorithm

(1) Initialization Centroid

Centroid is used as a central member of the cluster that expected to be an average of all the cluster members. If k cluster is formed, the number k centroid is necessary. There are various ways of doing initialization centroid, such as the random generation and the average spread. For example, interval data is [mindata$\cdots$maxdata]. Centroid random initialization is done with formula $C = \text{mindata} + (\text{maxdata} - \text{mindata}) \times \text{rand}\,()$. In this study, the initialization centroid for K-Means and FCM are used the average spread then the formula as follows:

$$\begin{aligned}
&\text{cluster number} \;=\; k \\
&s = (\text{maxdata} - \text{mindata})\,/k \\
&m = \text{mean}\,(\text{data}) \\
&C = [(m - s^*(k)) \;\; (\cdots) \;\; (m - s^*(1)) \;\; (m) \;\; (m + s^*(1)) \;\; (\cdots) \;\; (m + s^*(k))]
\end{aligned}$$

(2) Calculated Distance Neighborhoods

Neighborhoods distance is the distance between a centroid to each data that expressed by Eq. (2). According to [9] "an important step in most clustering is to select a distance measure, which will determine how the similarity of two elements is calculated". Thus, there are some varieties of distance function in clustering, including Euclidean distance (ED), Manhattan distance, Mahalanobis distance, and Hamming distance. In this study, K-Mean and FCM clustering using ED distance.

(3) Generated New Centroid
In this study, generating new centroid both of K-Means and FCM are based on the average all members cluster values. As an illustrated, for training data is as follows.

$$T = \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_5 \\ y_1 & y_2 & y_3 & y_4 & y_5 & y_6 \end{bmatrix}.$$

Next, the clustering results in an iteration is as follows.

$$G = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1\,0\,1\,1\,0\,0 \\ 0\,1\,0\,0\,1\,1 \end{bmatrix}.$$

Then, the new centroid is as follows.

$$c_1 = \frac{x_1 + x_3 + x_4}{y_1 + y_3 + y_4} \quad c_2 = \frac{x_2 + x_5 + x_6}{y_2 + y_5 + y_6}.$$

Meanwhile, in FCM, generating new centroid by using the COA (Center of Area) formula, Eq. (4).

2.3 Datasets

In this study, 152 days (from January–May 2016) daily network traffic usage of four client datasets from ICT unit were captured. Then, the data are analyzed by using MATLAB R2013b. The real dataset can be seen in Table 1 (Fig. 3).

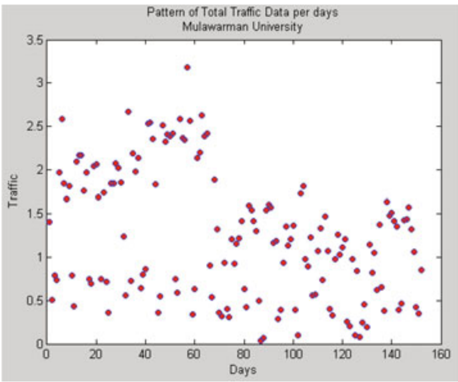


Fig. 3. Plot of network traffic per days

Table 1. The real of daily network traffic data (January–May 2016)

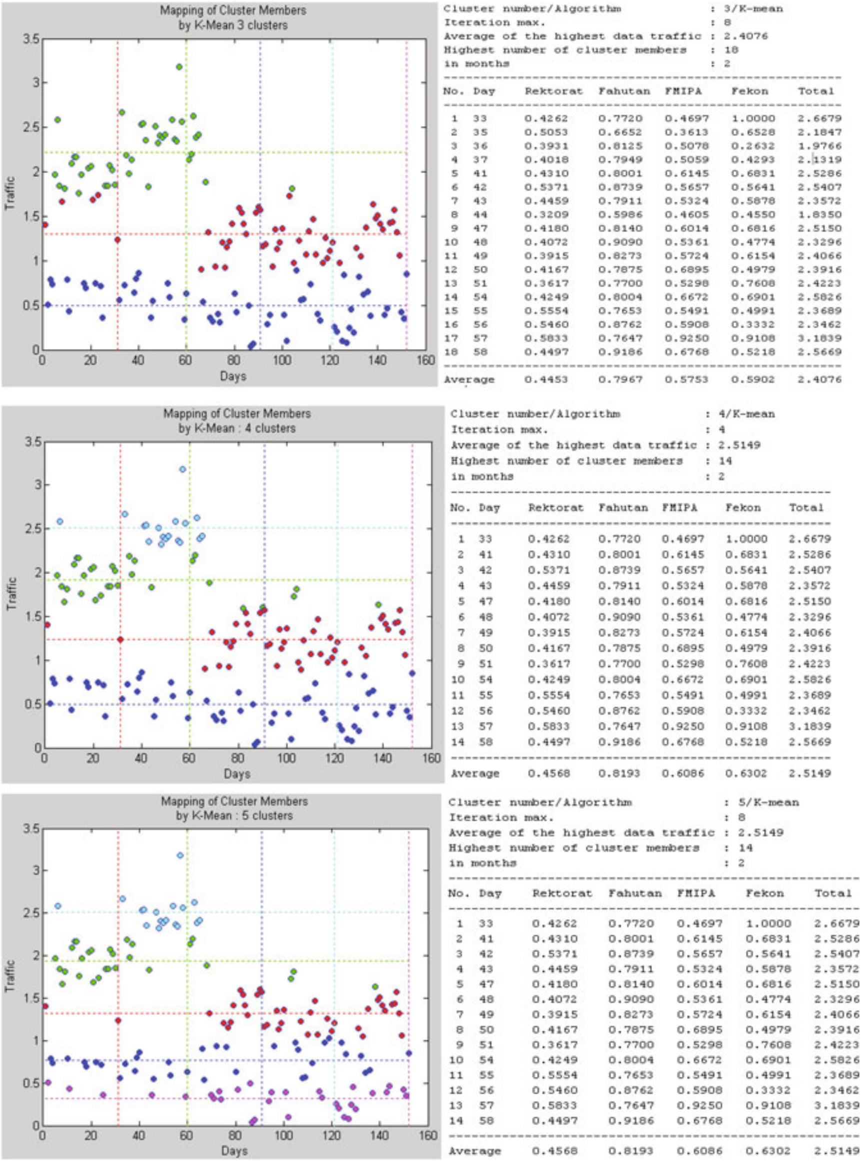
Rectorate	Forestry	Science	Economic	No	Rectorate	Forestry	Science	Economic
0.341	0.391	0.343	0.332	77	0.39	0.365	0.402	0
0.08	0.195	0.061	0.174	78	0.392	0.271	0.358	0.2
0.069	0.468	0.042	0.212	79	0.385	0.32	0.483	0.225
0.097	0.586	0.048	0.007	80	0.08	0.311	0.237	0.004
0.564	0.687	0.352	0.362	81	0.107	0.205	0.069	0.043
0.429	1	0.68	0.476	82	0.415	0.325	0.508	0.348
0.411	0.705	0.467	0.259	83	0.436	0.348	0.485	0.275
0.423	0.569	0.465	0.212	84	0.433	0.312	0.384	0.291
0.379	0.623	0.364	0.447	85	0.331	0.409	0.358	0.203
0.078	0.496	0.044	0.167	86	0.099	0.28	0.115	0.001
0.057	0.345	0.028	0.007	87	0.04	0	0	0.001
0.659	0.583	0.444	0.41	88	0.065	0	0	0.001
0.642	0.724	0.414	0.384	89	0.401	0.384	0.441	0.319
0.471	0.663	0.48	0.556	90	0.412	0.396	0.473	0.324
0.421	0.597	0.295	0.452	91	0.483	0.393	0.413	0.282
0.354	0.45	0.394	0.771	92	0.441	0.501	0	0.227
0.043	0.372	0.119	0.21	93	0.411	0.308	0.09	0.375
0.059	0.269	0.054	0.313	94	0.081	0.131	0.065	0.007
0.507	0.569	0.506	0.46	95	0.045	0.188	0.039	0.116
0.499	0.51	0.445	0.612	96	0.387	0.182	0.166	0.199
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
0.062	0.213	0.034	0	150	0.06	0.257	0.056	0.055
0.386	0.364	0.457	0.001	151	0.082	0.228	0.014	0.029
0.373	0.289	0.264	0	152	0.362	0.259	0	0.232

3 Results and Analysis

This section presents the empirical work and compares the experimental results of the K-Means and FCM algorithms on highest average usage network traffic problems. The performances are measured by the objective function value given by Eq. (1).

3.1 Analysis of K-Means

In this experiment, the scheme used is to classify the training data into 3, 4, and 5 cluster grouping patterns in order to observe the good cluster. The average data value is used as one of the centroid values. Then, for another centroid value is determined randomly in the space of data training. The results of K-Means are shown in Fig. 4.



Cluster number/Algorithm : 3/K-mean
Iteration max. : 8
Average of the highest data traffic : 2.4076
Highest number of cluster members : 18
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	35	0.5053	0.6652	0.3613	0.6528	2.1847
3	36	0.3931	0.8125	0.5078	0.2632	1.9766
4	37	0.4018	0.7949	0.5059	0.4293	2.1319
5	41	0.4310	0.8001	0.6145	0.6831	2.5286
6	42	0.5371	0.8739	0.5657	0.5641	2.5407
7	43	0.4459	0.7911	0.5324	0.5878	2.3572
8	44	0.3209	0.5986	0.4605	0.4550	1.8350
9	47	0.4180	0.8140	0.6014	0.6816	2.5150
10	48	0.4072	0.9090	0.5361	0.4774	2.3296
11	49	0.3915	0.8273	0.5724	0.6154	2.4066
12	50	0.4167	0.7875	0.6895	0.4979	2.3916
13	51	0.3617	0.7700	0.5298	0.7608	2.4223
14	54	0.4249	0.8004	0.6672	0.6901	2.5826
15	55	0.5554	0.7653	0.5491	0.4991	2.3689
16	56	0.5460	0.8762	0.5908	0.3332	2.3462
17	57	0.5833	0.7647	0.9250	0.9108	3.1839
18	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4453	0.7967	0.5753	0.5902	2.4076

Cluster number/Algorithm : 4/K-mean
Iteration max. : 4
Average of the highest data traffic : 2.5149
Highest number of cluster members : 14
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	41	0.4310	0.8001	0.6145	0.6831	2.5286
3	42	0.5371	0.8739	0.5657	0.5641	2.5407
4	43	0.4459	0.7911	0.5324	0.5878	2.3572
5	47	0.4180	0.8140	0.6014	0.6816	2.5150
6	48	0.4072	0.9090	0.5361	0.4774	2.3296
7	49	0.3915	0.8273	0.5724	0.6154	2.4066
8	50	0.4167	0.7875	0.6895	0.4979	2.3916
9	51	0.3617	0.7700	0.5298	0.7608	2.4223
10	54	0.4249	0.8004	0.6672	0.6901	2.5826
11	55	0.5554	0.7653	0.5491	0.4991	2.3689
12	56	0.5460	0.8762	0.5908	0.3332	2.3462
13	57	0.5833	0.7647	0.9250	0.9108	3.1839
14	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4568	0.8193	0.6086	0.6302	2.5149

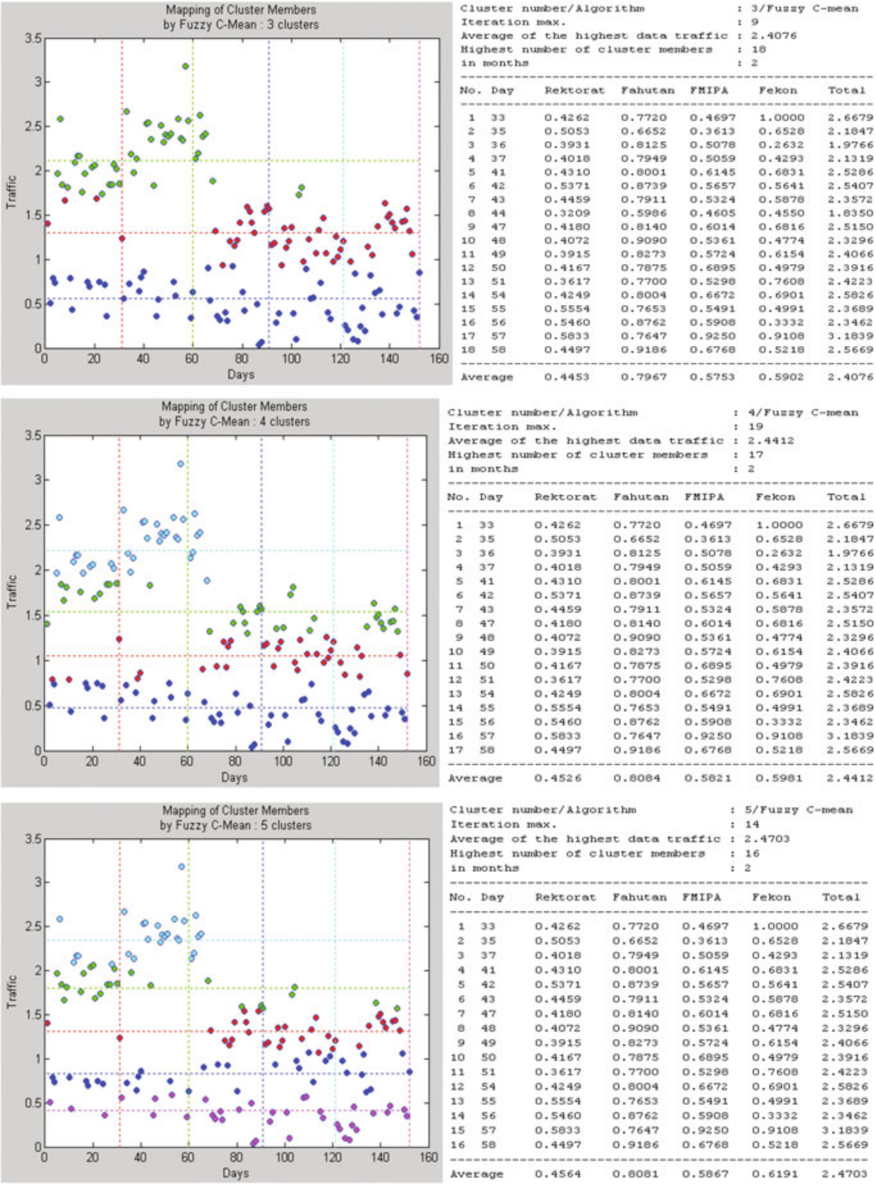
Cluster number/Algorithm : 5/K-mean
Iteration max. : 8
Average of the highest data traffic : 2.5149
Highest number of cluster members : 14
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	41	0.4310	0.8001	0.6145	0.6831	2.5286
3	42	0.5371	0.8739	0.5657	0.5641	2.5407
4	43	0.4459	0.7911	0.5324	0.5878	2.3572
5	47	0.4180	0.8140	0.6014	0.6816	2.5150
6	48	0.4072	0.9090	0.5361	0.4774	2.3296
7	49	0.3915	0.8273	0.5724	0.6154	2.4066
8	50	0.4167	0.7875	0.6895	0.4979	2.3916
9	51	0.3617	0.7700	0.5298	0.7608	2.4223
10	54	0.4249	0.8004	0.6672	0.6901	2.5826
11	55	0.5554	0.7653	0.5491	0.4991	2.3689
12	56	0.5460	0.8762	0.5908	0.3332	2.3462
13	57	0.5833	0.7647	0.9250	0.9108	3.1839
14	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4568	0.8193	0.6086	0.6302	2.5149

Fig. 4. Plot of clustering network traffic results using k-means

3.2 Analysis of Fuzzy C-Means

In this experiment, the scheme used is to classify the training data into 3, 4, and 5 cluster grouping patterns in order to observe the good cluster. The average data value is used as one of the centroid values. Then, for another centroid value is determined randomly in the space of data training. In this test, the FCM



Cluster number/Algorithm : 3/Fuzzy C-mean
Iteration max. : 9
Average of the highest data traffic : 2.4076
Highest number of cluster members : 18
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	35	0.5053	0.6652	0.3613	0.6528	2.1847
3	36	0.3931	0.8125	0.5078	0.2632	1.9766
4	37	0.4018	0.7949	0.5059	0.4293	2.1319
5	41	0.4310	0.8001	0.6145	0.6831	2.5286
6	42	0.5371	0.8739	0.5657	0.5641	2.5407
7	43	0.4459	0.7911	0.5324	0.5878	2.3572
8	44	0.3209	0.5986	0.4605	0.4550	1.8350
9	47	0.4180	0.8140	0.6014	0.6816	2.5150
10	48	0.4072	0.9090	0.5361	0.4774	2.3296
11	49	0.3915	0.8273	0.5724	0.6154	2.4066
12	50	0.4167	0.7875	0.6895	0.4979	2.3916
13	51	0.3617	0.7700	0.5298	0.7608	2.4223
14	54	0.4249	0.8004	0.6672	0.6901	2.5826
15	55	0.5554	0.7653	0.5491	0.4991	2.3689
16	56	0.5460	0.8762	0.5908	0.3332	2.3462
17	57	0.5833	0.7647	0.9250	0.9108	3.1839
18	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4453	0.7967	0.5753	0.5902	2.4076

Cluster number/Algorithm : 4/Fuzzy C-mean
Iteration max. : 19
Average of the highest data traffic : 2.4412
Highest number of cluster members : 17
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	35	0.5053	0.6652	0.3613	0.6528	2.1847
3	36	0.3931	0.8125	0.5078	0.2632	1.9766
4	37	0.4018	0.7949	0.5059	0.4293	2.1319
5	41	0.4310	0.8001	0.6145	0.6831	2.5286
6	42	0.5371	0.8739	0.5657	0.5641	2.5407
7	43	0.4459	0.7911	0.5324	0.5878	2.3572
8	47	0.4180	0.8140	0.6014	0.6816	2.5150
9	48	0.4072	0.9090	0.5361	0.4774	2.3296
10	49	0.3915	0.8273	0.5724	0.6154	2.4066
11	50	0.4167	0.7875	0.6895	0.4979	2.3916
12	51	0.3617	0.7700	0.5298	0.7608	2.4223
13	54	0.4249	0.8004	0.6672	0.6901	2.5826
14	55	0.5554	0.7653	0.5491	0.4991	2.3689
15	56	0.5460	0.8762	0.5908	0.3332	2.3462
16	57	0.5833	0.7647	0.9250	0.9108	3.1839
17	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4526	0.8084	0.5821	0.5981	2.4412

Cluster number/Algorithm : 5/Fuzzy C-mean
Iteration max. : 14
Average of the highest data traffic : 2.4703
Highest number of cluster members : 16
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	35	0.5053	0.6652	0.3613	0.6528	2.1847
3	37	0.4018	0.7949	0.5059	0.4293	2.1319
4	41	0.4310	0.8001	0.6145	0.6831	2.5286
5	42	0.5371	0.8739	0.5657	0.5641	2.5407
6	43	0.4459	0.7911	0.5324	0.5878	2.3572
7	47	0.4180	0.8140	0.6014	0.6816	2.5150
8	48	0.4072	0.9090	0.5361	0.4774	2.3296
9	49	0.3915	0.8273	0.5724	0.6154	2.4066
10	50	0.4167	0.7875	0.6895	0.4979	2.3916
11	51	0.3617	0.7700	0.5298	0.7608	2.4223
12	54	0.4249	0.8004	0.6672	0.6901	2.5826
13	55	0.5554	0.7653	0.5491	0.4991	2.3689
14	56	0.5460	0.8762	0.5908	0.3332	2.3462
15	57	0.5833	0.7647	0.9250	0.9108	3.1839
16	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4564	0.8081	0.5867	0.6191	2.4703

Fig. 5. Plot of clustering network traffic results using fuzzy C-means

initial parameter values were implemented, (a). Number of clusters 3, 4, and 5; (b). Maximum iterations = 50; and (c). Stopping Criteria (ξ) = 10^{-5}. In this research, applied testing scheme is the amount of usage per day of data network traffic from the four parts (rectorate, forestry, science, and economic). Then, the

Table 2. Results comparison K-means and FCM methods

Cluster	Methods	Highest cluster members number	Average of network traffic usage	Highest of network traffic usage month
3	KCM	18	2.4076	2
	FCM	18	2.4076	2
4	KCM	14	2.5149	2
	FCM	17	2.4412	2
5	KCM	14	2.5149	2
	FCM	16	2.4703	2

highest, middle and lowest data values are obtained. The results of FCM are shown in Fig. 5.

In this study, the accuracy level in order to get the centroid value by using FCM method is quite careful than K-Means method, Table 2.

4 Conclusion

In this paper we presented a comparison of K-Means and FCM methods then compared the centroid accuracy by using various performance criteria. This research was used network traffic usage from four units; rectorate, forestry, science, and economic. In clustering, the examining of 1 parameter of centroid value is 3 centroid values. Our experiment showed that the FCM method is better results analysis in clustering than K-Means. Nevertheless, we have also concluded that FCM algorithm was slower than K-Means. As future work, an optimizing methods in order to get good accuracy between centroids is proposed.

References

1. Arroyo A, Herrero A et al (2016) Analysis of meteorological conditions in spain by means of clustering techniques. J Appl Logic. doi:10.1016/j.jal.2016.11.026.
2. Bai C, Dhavale D, Sarkis J (2016) Complex investment decisions using rough set and fuzzy c-means: an example of investment in green supply chains. Eur J Oper Res 248(2):507–521
3. Carvalho L, Barbon S Jr et al (2016) Unsupervised learning clustering and self-organized agents applied to help network management. Expert Syst Appl 54(C):29–47
4. Elaal A, Hefny H, Elwahab A (2010) Constructing fuzzy time series model based on fuzzy clustering for a forecasting. J Comput Sci 6(7):735–739
5. Huang CK, Hsu WH, Chen YL (2013) Conjecturable knowledge discovery: a fuzzy clustering approach. Fuzzy Sets Syst 221(2):1–23
6. Pandit YP, Badhe YP et al (2011) Classification of Indian power coals using k-means clustering and self organizing map neural network. Fuel 90(1):339–347

7. Stetco A, Zeng XJ, Keane J (2015) Fuzzy c-means++: fuzzy c-means with effective seeding initialization. Expert Syst Appl 40(21):7541–7548
8. Sucasas V, Radwan A et al (2016) A survey on clustering techniques for cooperative wireless networks. Ad Hoc Netw 47(C):53–81
9. Velmurugan T (2014) Performance based analysis between k-means and fuzzy c-means clustering algorithms for connection oriented telecommunication data. Appl Soft Comput 19(6):134–146

RDEU Evolutionary Game Model and Simulation of the Network Group Events with Emotional Factors

Guoqiang Xiong(✉), Xian Wang, Ying Yang, and Yuxi Liu

School of Business Administration, Xi'an University of Technology,
Xi'an 710054, People's Republic of China
xgq168@163.com

Abstract. At present, network communication becomes more efficient and geographical restrictions barely exist. However, the network mass incidents tend to be multiple and frequent, which are mostly unexpected. In this paper, we point at studying the limitations of the traditional game theory in the players' rational assumptions, and consider the influence of the irrational factors on the netizens behavior. Based on RDEU theory and evolutionary game theory, a RDEU evolutionary game model of group events on network is constructed. In addition, we studied the evolution mechanism of game Nash equilibrium under different emotional conditions and the numerical simulation of the model, which is carried out by using the numerical simulation method of Matlab. After that, the evolutionary equilibrium state of the game between the netizens in different emotional states of the network group events is analyzed. Finally, research results show that emotional factors have a great impact on the network behavior of group events; when the emotions of netizens are pessimistic, the game participants are more inclined to confrontational behavior, and will have both the conflicting herd effect, more prone to network mass incidents. In this case, it is more likely to produce group events on network.

Keywords: Network group event · RDEU evolutionary game model · Emotion · Simulation

1 Introduction

Generally, the whole Society is in the period of social transformation filled with various contradictions and social conflicts, including the increasing number of network mass incidents and the negative influence from that. We often find a simple network information will attract tens of millions of click rate in a short period, an infectious emotion may spread out of control like a virus, thus causing emotional excitement. That is to say, the development of the group events on networks is often due to the interaction between social and human factors.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_25

The evolution of this interaction has the derivative harmfulness, easily leading to systemic economic and social crisis. Thus, social stability is facing a severe challenge. Therefore, keeping in touch with the emotional state of netizens in time is of great significance to grasp the law of the evolution in the behavior of netizens group and monitor the occurrence of early warning derivative events.

At present, a majority of researchers of the network mass incidents focused on the aspects of the network conflict. Morales [5] believed that network mass incidents refer to the large-scale pooling of netizens' opinions in the network for specific hotspot problems, thus affecting the real-life group events. Levine [2] argued that cyberspace disputes and conflicts, which expand the interaction of people, can activate the public awareness of the society in order to create a new public realm, and provide new environmental resources for democracy, effectively avoid autocratic. In addition, there are some research focus on the function of network conflict, the conflict on real society and democratic system. Post [6] pointed out that the traditional legal control tools are confronting severe network conflict challenges. Technicians cannot depends on the traditional management policy to effectively restrict network environment and reconstruct the rules' system of cyberspace. Qiu [7] summarized three types of network mass incidents: First, some incidents occur in the Internet but exerts substantial influence on the real society; the second type of incidents happens in the reality with the virtual online organization as the intermediary; the third one happens through offline interaction that is based on the online communication. Some scholars proposed corresponding strategies to solve the problem about unexpected group incidents from different perspectives. From the view of communication studies, Zhu et al. [13] emphasized that the fundamental effect of faith on the public opinion and the guidance of the "opinion leader" for the direction of public opinion. From the social and psychological point of view to investigate the social struggle, Sidney et al. [10] argued it was the existence of a sense of deprivation of the individual that resulted in resentment, thus forming a collective behavior. It follows that the present scholars' research on network unexpected group incidents mostly applies the qualitative, canonical approaches and comes from the single perspective but seldom establishes mathematical models based on the theories of management science to make relevant research. At present, a few scholars as Liu [3,4], Deng et al. [1], Xie [12], Wang [11] and Sun [9] made research on traditional unexpected group incidents. Aiming at the conflict evolutionary mechanism of unexpected group incidents, they established the models of unexpected group incidents and initially revealed the mechanism of such incidents. Although evolutionary game models of unexpected group incidents deepened the research concerning the conflict evolutionary mechanism by dividing up the social groups into advantaged groups and vulnerable groups, they did not consider the influence of emotional factor on the state of evolutionary game equilibrium in decision-making in the game.

All in all, the study of the evolution mechanism of the group events on networks is still in its infancy, and only qualitative methods have been used to analyze the causes, types and countermeasures of the events. There are little research on the evolution mechanism of the group events on networks, especially we lack the direct research on game model of endogenous factors like emotion to further explore the evolution of the event. Based on the Quiggin's RDEU theory (Rank-dependent expected utility) [8], this paper constructs the RDEU evolutionary game model of the group events on networks and studies the evolvement process of netizens' strategy selection under different emotional states, and provides emergency decision support for policy selection of the group events on networks. There are four sections in this paper. In the second section, we introduce the emotional function which reflects the psychological activities of netizens and establish the RDEU game model of the behavior mechanism of the group events on networks. The third section, according to the different emotions of netizens, adopts Matlab's numerical simulation analysis of evolution equilibrium on the netizens RDEU game model under six evolutionary scenarios (for example, netizens of one party are rational, while the other party is pessimistic, netizens are both pessimistic, etc.). The fourth section is the summary of the full text.

2 The RDEU Evolutionary Game Model among Netizen Group

2.1 The Method of Constructing the Netizen Game Model

In the network unexpected group incident, the strategy sets of the netizen groups can be divided into: "fight" (represented by F) and "peace" (indicated by P), the former refers to radical articles or comments while the latter refers to quiet observation of the course of events without taking any radical actions. In that game, there are three different situations: (1) When both of the two groups of netizens adopt the sheep-flock strategy to "F"; the influence of the incident may be accordingly expanded, and both of interactive objects may respectively gain interaction revenues represented by "V"; extraneous revenues is implied by "S"; paying the costs is implied by "C". Then, both sides should pay the costs and make extraneous and interaction revenues; (2) When two netizen groups adopt different strategies, the principal party of "P" can make interaction revenues without paying "C" while the principal party of "F" pays "C" and makes "V" and "S"; (3) When both groups apply the peaceful sheep-flock strategy, namely, as onlookers or passers-by, neither of them gain any revenue or deliver any contribution to the incidents, make any "S" or pay any costs, so they spend no cost and gain no benefit. Table 1 shows the revenue matrix in the game of the netizen's sheep-flock effect in the network unexpected group incidents.

Table 1. Revenue matrix in the game of the netizen

		Netizen B	
		Fight (F) q	Peace (P) $1-q$
Netizen A	Fight (F) p	$2V+S-C, 2V+S-C$	$V+S-C, V$
	Peace (P) $1-p$	$V, V+S-C$	0, 0

Table 1 shows that the netizens pay a little price when they forward or comment the article in most case. To be concluded, the fight brings more benefits than peace, obviously $S > C$, so the size relation between parameters is: $2V + S - C > V + S - C > V > 0$.

2.2 Establishing the RDEU-Based Evolutionary Game Model for the Netizen Groups

The dynamic game model was basically built on the basis that the netizens were entirely rational. However, when the social economic environment and the decision problems are more complicated, people's rationality will be obviously limited. For this reason, in a move to ensure the application value of game analysis, the netizen game of limited rationality must be analyzed. Consequently, the following part will discuss the evolutionary game model among netizen groups in a drive to analyze the game under the analytical framework of repeated random-pair game among netizens of low rationality. Suppose that the netizens adopting "F" occupy the proportion of "p" of the netizen group A and the counterparts adopting "P" account for $(1 - p)$; the netizens adopting "F" take up a proportion of "q" of the netizen group B and the counterparts adopting "P" account for $(1 - q)$, where $p, q \in [0, 1]$.

Network unexpected group incidents have displayed distinctive characteristics from traditional group incidents chiefly in terms of emotional extremity in these years. Because of lack of authoritative review mechanism in communication, some features, such as popularization and randomness arose. It is correspondingly difficult for the governmental institutions to control. At the same time, because of some problems like enlarging gap of wealth, soaring house prices and inappropriate governmental behaviors, the netizens tend to show sympathy to the disadvantaged groups. When the incident object relates to social unfairness, the netizens are evidently easier to get agitated or go to extremes. To reflect the emotional status of both parties of the game in strategy selection, the paper hereby applies the RDEU theory to include the emotional factors of the netizen groups into the above game model of the netizen's sheep-flock effect. Suppose the emotional function of netizen group A and netizen group B is respectively $w_A(p) = p^{r_1}$ and $w_B(q) = q^{r_2}$ and $r_1, r_2 > 0$, called the emotional index of the netizens respectively.

In the RDEU theory, we can get that the probabilities of different revenues made by netizen group A and B when they adopt different strategies and the rank of the revenue value of different strategy portfolio among all the revenue values. It can be called the rank of corresponding revenue. If the rank of a strategy portfolio is higher, it is less possibility that the revenues of other strategy portfolios can exceed it. That is the implication of "high revenue and low probability". The corresponding decision weights are calculated with probabilities and ranks of different revenues, as listed in Tables 2 and 3.

Table 2. The probability distribution, rank and decision weights of netizen group A's revenue

The revenue of Netizen group A x_i	$2V+S-C$ (F)	$V+S-C$ (F)	V (P)	0 (P)
Probability p_i	pq	$p(1-q)$	$(1-p)q$	$(1-p)(1-q)$
Rank RP_i	1	$1-pq$	$1-p$	$(1-p)(1-q)$
Decision weights $\pi(x_i)$	$w_A(pq)$	$w_A(p)-w_A(pq)$	$w_A(p+q-pq)-w_A(p)$	$1-w_A(p+q+pq)$

Table 3. The probability distribution, rank and decision weights of netizen group B's revenue

The revenue of Netizen group A x_i	$2V+S-C$ (F)	$V+S-C$ (F)	V (P)	0 (P)
Probability p_i	pq	$(1-p)q$	$p(1-q)$	$(1-p)(1-q)$
Rank RP_i	1	$1-pq$	$1-q$	$(1-p)(1-q)$
Decision weights $\pi(x_i)$	$w_B(pq)$	$w_B(q)-w_B(pq)$	$w_B(p+q-pq)-w_B(q)$	$1-w_B(p+q+pq)$

U_{AF} and U_{AP} represent the expected revenues of netizen group A adopting "*F*" and "*P*" respectively in the network unexpected group incidents.

$$U_{AF} = V(pq)^{r_1} + (V+S-C)p^{r_1},$$
$$U_{AP} = V(p+q-pq)^{r_1} - Vp^{r_1}.$$

The RDEU expected revenues of netizen group A as:

$$U_A(p,q;w) = U_{AF} + U_{AP} = V(pq)^{r_1} + V(p+q-pq)^{r_1} + (S-C)p^{r_1}. \quad (1)$$

Similarly, U_{BF} and U_{BP} are taken to represent the expected revenues of netizen group B adopting "*F*" and "*P*"

$$U_{BF} = V(pq)^{r_2} + (V+S-C)q^{r_2},$$
$$U_{BP} = V(p+q-pq)^{r_2} - Vq^{r_2}.$$

The RDEU expected revenues of netizen group B as:

$$U_B(p,q;w) = U_{BF} + U_{BP} = V(pq)^{r_2} + V(p+q-pq)^{r_2} + (S-C)p^{r_2}. \quad (2)$$

$\overline{U_A}$ is taken to represent the average revenue of netizen group A, then

$$\begin{aligned}\overline{U_A} =& p^{r_1}U_{AF} + (1-p^{r_1})U_{AP} = Vp^{2r_1}q^{r_1} - Vp^{r_1}(p+q-pq)^{r_1} \\ &+ (2V+S-C)p^{2r_1} + V(p+q-pq)^{r_1} - Vp^{r_1}. \quad (3)\end{aligned}$$

The limited rationality of the netizen groups indicates that people will find better strategies by learning in the game instead of finding the best strategy at the beginning. According to the theory of replicated dynamics for biological evolution, the game player who adopts strategies yielding lower revenues will change the strategy and turn to (simulate) the strategies yielding higher revenues. For this reason, the proportion of group members who apply different strategies may change, then the speed of proportion change and the proportion of specific strategies have positive relation with the margin of its revenue over the average revenue. Consequently, the speed of change of "p" of the netizens resorting to "F" among the above netizen group A can be shown in the replicated dynamic equation as follows:

$$\begin{aligned}\frac{dp}{dt} &= p^{r_1}(U_{AF} - \overline{U_A}) = p^{r_1}(1 - p^{r_1})[V(pq)^{r_1} \\ &\quad + (2V + S - C)p^{r_1} - V(p + q - pq)^{r_1}].\end{aligned} \tag{4}$$

At the same time, is taken to represent the average revenue of netizen group B, so

$$\begin{aligned}\overline{U_B} &= q^{r_2}U_{BF} + (1 - q^{r_2})U_{BP} = Vp^{r_2}q^{2r_2} - Vq^{r_2}(p + q - pq)^{r_2} \\ &\quad + (2V + S - C)q^{2r_2} + V(p + q - pq)^{r_2} - Vq^{r_2}.\end{aligned} \tag{5}$$

When the netizen group B take the "fight" behavior, the change of speed of netizen number's proportion can be represented as the below replicator dynamics equation as

$$\begin{aligned}\frac{dq}{dt} &= q^{r_2}(U_{BF} - \overline{U_B}) = q^{r_2}(1 - q^{r_2})[V(pq)^{r_2} \\ &\quad + (2V + S - C)q^{r_2} - V(p + q - pq)^{r_2}].\end{aligned} \tag{6}$$

Simultaneous integration of the replicated dynamic equations of netizen group A and B may help to build the replicated dynamic model for the evolutionary game of netizen group A and B applying changing strategies as time changes.

$$\begin{cases}\frac{dp}{dt} = p^{r_1}(1 - p^{r_1})[V(pq)^{r_1} + (2V + S - C)p^{r_1} - V(p + q - pq)^{r_1}] \\ \frac{dq}{dt} = q^{r_2}(1 - q^{r_2})[V(pq)^{r_2} + (2V + S - C)q^{r_2} - V(p + q - pq)^{r_2}].\end{cases} \tag{7}$$

3 Numerical Simulation of RDEU-Based Evolutionary Game Model

Because of the numerous parameters in Formula 7 and the uncertainty of the parameter values, this part adopts the approach of numerical simulation to further discuss the afore-mentioned model in order to reflect how emotions influence the behaviors of netizens. To facilitate the operation, we suppose V = 6, S = 8 and C = 4 and rewrite the replicated dynamic model into:

$$\begin{cases}\frac{dp}{dt} = p^{r_1}(1 - p^{r_1})[6(pq)^{r_1} + 16p^{r_1} - 6(p + q - pq)^{r_1}] \\ \frac{dq}{dt} = q^{r_2}(1 - q^{r_2})[6(pq)^{r_2} + 16q^{r_2} - 6(p + q - pq)^{r_2}].\end{cases} \tag{8}$$

After that we assume r_1 and r_2 which the emotional indexes of the netizens be divided into six kinds of conditions, then the conditions simulation software named matlab7.0 will be applied to simulate the changes of the strategies of the network groups in network unexpected group incidents and analyze the different conditions to reach the state of evolutionary equilibrium. Assuming the initial ratio of p and q is (0.2, 0.7). When t = 0, the proportion of netizens of group B applying "*F*" is higher than that of group A.

Situation 1: Both sides of netizen groups are rational, namely $r_1 = r_2 = 1$.

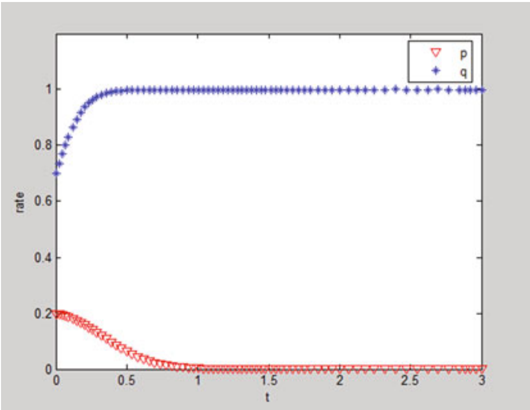


Fig. 1. The result of simulation when $r_1 = 1$, $r_2 = 1$

Figure 1 shows that the evolutionary result of the experiment when all the parameters conform to situation 1. The result manifests that when both the netizen group A and netizen group B are rational, the system will eventually evolve to be "*F*" of group B and "*P*" of group A.

Situation 2: One of the netizen groups is rational while the other is optimistic.

$r_1 = 1$ and $r_2 \in (1, +\infty)$ or $r_1 \in (1, +\infty)$ and $r_2 = 1$, suppose $r_1 = 1, r_2 = 1.7$ or $r_1 = 1.7, r_2 = 1$

Figure 2 describes the evolutionary results of the system when all the parameters satisfy to situation 2. The result shows that when netizen group A and group B are optimistic or group A are optimistic and group B are rational, the system will finally evolve to be "*F*" of Group B and "*P*" of Group A.

Situation 3: One party of the netizen groups is rational while the other is pessimistic. $r_1 = 1$ and $r_2 \in (0, 1)$ or $r_1 \in (0, 1)$ and $r_2 = 1$, suppose $r_1 = 1, r_2 = 0.9$ or $r_1 = 0.9, r_2 = 1$.

Figures 3(a) and (b) indicate that the evolutionary results of the system when all the parameters conform to situation 3. The result shows that when netizen group A are rational and group B are pessimistic or group A are pessimistic and group B are rational, the system will finally evolve to be "*F*" of group B and "*P*" of group A.

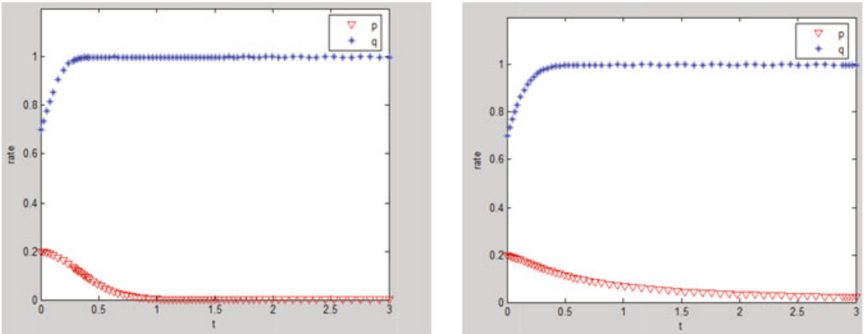


Fig. 2. (a) The result of simulation when $r_1 = 1$, $r_2 = 1.7$; (b) The result of simulation when $r_1 = 1.7$, $r_2 = 1$

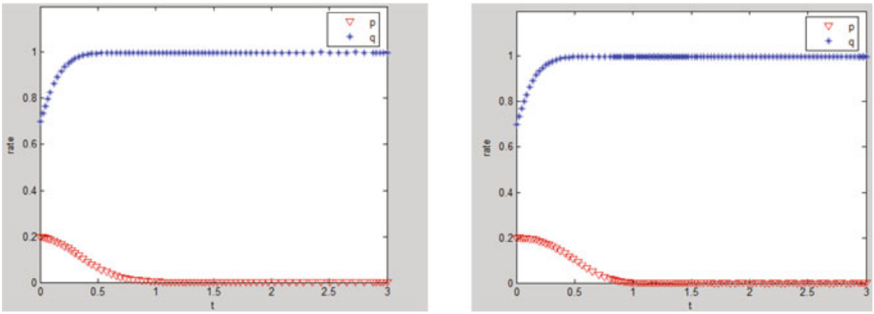


Fig. 3. (a) The result of simulation when $r_1 = 1$, $r_2 = 0.9$; (b) The result of simulation when $r_1 = 0.9$, $r_2 = 1$

Situation 4: Both sides of netizen groups are pessimistic, as $r_1 \in (0, 1)$ and $r_2 \in (0, 1)$, suppose $r_1 = 0.5, r_2 = 0.6$.

Figure 4 describes that the evolutionary result of the system when all the parameters conform to situation 4. The result shows that when netizen group A and B are both pessimistic, the system will finally evolve to generate the sheep-flock effect that both sides apply "F" because netizen group A simulates netizen group B.

Situation 5: Both sides of netizen groups are optimistic, as $r_1 \in (1, +\infty)$ and $r_2 \in (1, +\infty)$, suppose $r_1 = 1.2, r_2 = 1.3$.

Figure 5 depicts the evolutionary result of the system when all the parameters conform to situation 5. The result shows that when netizen group A and B are both optimistic, the system will finally evolve to be the "F" of netizen group B and the "P" of netizen group A.

Situation 6: One party of the netizen groups is optimistic while the other is pessimistic, namely, $r_1 \in (0, 1)\, and\, r_2 \in (1, +\infty)$ or $r_1 \in (1, +\infty)\, and\, r_2 \in (0, 1)$, suppose $r_1 = 0.9, r_2 = 1.1$ or $r_1 = 1.1, r_2 = 0.9$

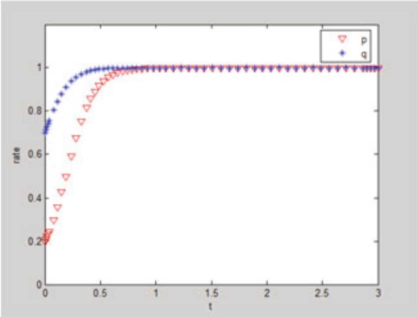


Fig. 4. The result of simulation when $r_1 = 0.5$, $r_2 = 0.6$

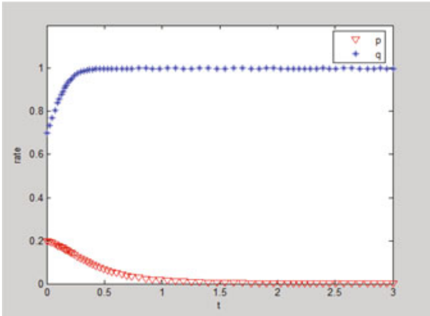


Fig. 5. The result of simulation when $r_1 = 1.2$, $r_2 = 1.3$

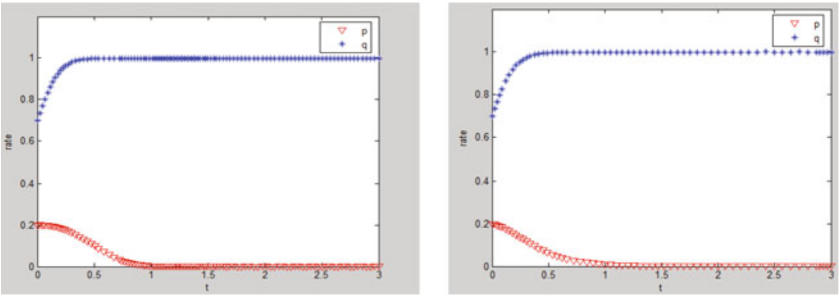


Fig. 6. (a) The result of simulation when $r_1 = 0.9$, $r_2 = 1.1$; (b) The result of simulation when $r_1 = 1.1$, $r_2 = 0.9$

Figure 6(a) and (b) describe the evolutionary results of the system when all the parameters conform to situation 6. The result shows that when netizen group A and group B are optimistic or group A and group B are pessimistic, the system will finally evolve to be of Group B and of group A.

4 Conclusion

Based on the shortcomings of the existing literature on the research of group events on networks and the new RDEU theory, this paper establishes RDEU game model of group events on network. Under the six evolutionary scenarios, it analyzes the extent and condition of influence of different emotions on netizens' game behavior. This paper overcomes the limitations of the traditional game model which on the basis of "full rationality" hypothesis of netizens, and considers the influence and mechanism of endogenous factors such as netizen emotion on the evolution of the group events on networks. It also reveals the deeper root of the outbreak and expansion of group events on networks. The results show that emotional factors have an important influence on the choice of the game strategies of the group events on networks. When some netizens has a “pessimistic” emotion, they tend to “antagonistic” behavior, and the higher the degree of pessimism, the more likely they choice a risk strategy; When some netizens has a “optimistic” emotion, they are easy to commit “concessions” behavior; When the netizens are pessimistic in both groups, the event is likely to evolve into a conflict effect with the herd effect.

The conclusion of this study can provide some reference and inspiration for the emergency management and prevention and control measures of network mass incidents. Government departments or social media should increase the response mechanism for the interests and wishes of netizens, strengthen netizens’ emotion monitoring and analysis and provide psychological intervention and emotional counseling for them timely.

References

1. Deng Y, O’Brien KJ (2012) Relational repression in China: using social ties to demobilize protesters. Soc Sci Electron Publishing 215(215):533–552
2. Levine P (2001) Civic renewal and the commons of cyberspace. Nat Civic Rev 90(3):205–212
3. Liu D, Wang W, Li H (2013) Evolutionary mechanism and information supervision of public opinions in internet emergency. Procedia Comput Sci 17:973–980
4. Liu D, Han C, Yin L (2016) Multi-scenario evolutionary game analysis of evolutionary mechanism in urban demolition mass incident. Oper Res Manage Sci 1:76–84
5. Morales AJ, Borondo J et al (2014) Efficiency of human activity on information spreading on twitter. Soc Netw 39(1):1–11
6. Post DG (2002) Against ’against cyberanarchy’. SSRN Electron J 17(4):1365–1387
7. Qiu W (2009) The study of the network group events. J Ningbo Radio TV Univ 3:1–4 (in Chinese)
8. Quiggin J (1982) A theory of anticipated utility. J Economic Behav Organ 3(4):323–343
9. Sun H, Wang X, Xue Y (2016) Stochastic evolutionary game model for unexpected incidents involving mass participation based on different scenarios. Oper Res Manage Sci 4:23–30 (in Chinese)
10. Tarrow SG (1994) Power in movement: social movements and contentious politics. Contemp Sociol 28:3–12
11. Wang X, Li Y, Sun H (2015) Evolutionary game analysis of unexpected incidents involving mass participation based scenario inference. J Manage Sci 6:133–143 (in Chinese)
12. Xie B, Zhang W et al (2014) Evolutionary game analysis and application to mass emergency under government coordination. Operations Research and Management Science 5:243–249 (in Chinese)
13. Zhu J, Zhang E, Cui Y (2013) Characteristics and causes of the network group events. Theory Res 32:96–97 (in Chinese)

Effects of Internet Word-of-Mouth of a Tourism Destination on Consumer Purchase Intention: Based on Temporal Distance and Social Distance

Mo Chen(✉) and Jingdong Chen

Faculty of Economy and Management, Xi'an University of Technology, Xi'an 710054, People's Republic of China
chenmo19921201@foxmail.com

Abstract. The proportion of tourists who referencing the Internet word-of-mouth increased. Hence, tourism destinations' online word-of-mouth has an practically significant effect on the willing of purchase. Due to an increasing trend about the proportion of tourists referencing Internet word-of-mouth (IWOM) for, making a study of tourism destination online word-of-mouth influence on purchase intention has practical significance. The purpose of this article is to provide marketing advises in different temporal and social distance. The research methods of this article is dividing 110 college students into two groups. All the students are allowed to read online word-of-mouth about two destinations, and this is for measuring consumers' tourism consumption which under different temporal distance and social distance categories. The conclusion is that quality of word-of-mouth exercise a greater effect for the consumers when they will travel in the far future. And the quantity of word-of-mouth influence nearly future purchase intention of consumers more prominently. For far Social distance between reader and IWOM witter, purchase intention is more significantly affected by the quality of word-of-mouth. On the contrary, purchase intention is more significantly influenced by the quantity word-of-mouth, when they feel they are close.

Keywords: Internet · Word-of-Mouth (IWOM) · Consumers purchase intention · Temporal distance · Social distance

1 Introduction

The context of traveling getting increasingly popular, prospects is taking on prosperity and potential is substantial for the traveling sector in China, with online consumption in doing sightseeing projected to reaching RMB 800 Billion. In 2014, this country had a transaction scale of RMB 3250 Billion, in which the E-market of tourism represents 11.3% to 367 Billion, a 45.6% rise compared to last year.

The reason why Chinese marketing has good developing prospects is great influence of WOM. As the consequence surveyed by Market research company

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_26

GfK shows that more than half of Chinese travelers rely on personal contact to obtain travel information, including friends, family and colleagues, which is same as online reading. In terms of actual booking, online instead of offline, there are 63% of visitors choosing online payment, which means that they are unavoidable to read some online information about destinations. 80% of consumers said they buy, influenced by the Internet word-of-mouth IWOM, implying that researching the tourism destination IWOM is of practical significance. At the same time, customer get tourism information platforms and ways, which are more diverse, because the Internet and mobile phone mobile client widely are used. The IWOM regarding tourism destinations in platforms such as ctrip, weibo or other tourism BBS or APP spread more widely, showing it exerts a great influence.

On the one hand, because the time intervals from customers reading IWOM to the actual purchase is different, leading to their purchase will be different. Customers will have a psychological judgment when they read online word-of-mouth. The judgment of the psychological distance between customers and the author of IWOM will also change the customer's purchase intention. Therefore, this paper starts from the difference of the influence of the word-of-mouth in the different temporal distance and social distance, observes the performance of consumers' purchasing intention and explains the influence of online word-of-mouth.

2 Literature

2.1 Word-of-Mouth Research

Arndt [1] first raised word-of-mouth in marketing research significance and showed word-of-mouth as a third party's influences on consumer behavior. And early studies tended to probe into the reason for word of mouth research [7]. With the development of the Internet and the increase of its scale, word-of-mouth research started in the Internet-word-of-mouth. In research, Internet-word-of-mouth (IWOM) [24], electronic word-of-mouth (ewom) [22], and electronic reputation refer to the same concept. In the study of IWOM, scholars [6,13] thought the IWOM is that customers or potential customers in the network community or BBS put forward for the product or business the positive and negative. Dong dahai defined the concept of IWOM [8], and considered IWOM has three main differences compared to traditional word-of-mouth: ① Spread more broadly. ② Connection between the receivers and senders is not determined [4]. In addition to the strong relationship between circles of friends, there is also a weak relationship. ③ commercial uncertainty. Due to some reasons such as the anonymity of the Internet behavior, Internet word of mouth will be controlled by some enterprises or business organizations. So, examining IWOM is different from traditional marketing and this makes word-of-mouth a fresh meaning.

From the beginning of the 21st century's China, research on word-of-mouth started to be increasing gradually. Domestic study about word-of-mouth for consumers to buy has two directions. One direction is discovering the emotional

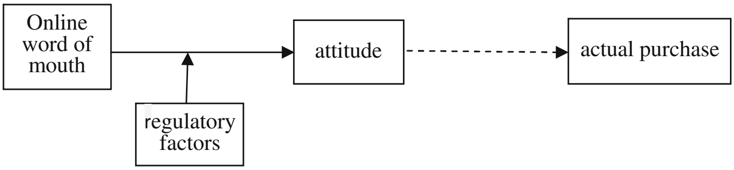


Fig. 1. Word-of-mouth and attitude theoretical model

factors of word-of-mouth. Guo [10] put forward the theoretical model of word-of-mouth for consumer attitudes, as shown in Fig. 1. After the model raised, A lot of research around the positive word-of-mouth and negative word-of-mouth. They believe that the emotion behind IWOM influence actual purchase through attitude. Baber et al. [2] found that the positive online word-of-mouth can bring positive change on consumers' purchase intentions. Bailey [3] proved that negative reviews have a significant impact on customers decision. Based on the original model studied, concentrated on regulating variables, which joined the adjustment factors such as trust and relationship strength. Balaji [22] found that the relationships among language divergence effect positive word-of-mouth intentions.

On the other hand, Zuo [20] put forward a model concerning the consumer purchase intention affected by word-of-mouth's quantity and quality. The existing literature has proved quality and quantity of word-of-mouth influencing intentions of consumers purchase [23]. This studies focus on the transmission of information. Balajia [4] mentioned information quality in the article about language divergence in WOM. This article in also refer to interaction quality and relationship quality.

2.2 Temporal Distance and Social Distance

Research on consumers' behavior and psychology research on time distance and social distance in recent years also have attracted more attention. With the development of the theory of psychological explanation level, researchers began to define the temporal distance and social distance from the Angle of cognitive. Reference point means the subject is located in someplace at the moment, the distance between the location of reference point and object is vague [19]. Time distant refers to people's feeling about the interval from an event occurrence to now [5], while social distant is they feeling about the relationship between relevant person and them [5].

Many studies probe into social distance and time distance with purchase intention of literature. Junghye's [15] conclusions are drawn that with the time interval between intention formation and the action becoming longer, and the greater the spatial distance to a destination is, the non-travelers have higher the probability to change behaviors. Xu [14], through experiment to measure the temporal distance and social distance for the interaction of the virtues and vices products, proved discrepancy in customers' preferences for the product under different situations. Zhang [21] proved the temporal distance effect on consumers' impulse consumption. Zou [25] found that the customers emphasis desirability or feasibility differently, when the consumer choose financial products under different temporal distances.

The preliminary study shows that temporal and social distance have an effect on consumer purchase intention. In short, combined with psychology and word of mouth is also a relatively new research.

3 The Experimental Research

3.1 Research Hypothesis

Internet word-of-mouth has one of the biggest differences compared with the traditional word of mouth, whose access is of diversification, so the quantity of electronic word-of-mouth is substantial. In terms of tourism destination's IWOM, consumers can collect and obtain, from several ways like tourism network BBS, blog, Micro-blog, instant messaging (We-chat, qq), network video, Wikipedia and travel quiz, tourist IWOM. Liu [16] proved that the movie aspect through word-of-mouth affects the relationship between the number of product sales and reputation. The study of the quantity of word-of-mouth, with the development of the network, from the original single-platform to cross-platform research, at present, has developed [11] swiftly, so the total quantity of IWOM is measured from the viewpoint of cross-platform.

The quality of word-of-mouth in essence refers to the content of word of mouth. In terms of tourist destination of IWOM, word-of-mouth quality evaluation is mainly from the correlation of word of mouth to convey information, professional degrees, and the vitality of comprehensive evaluation on the quality of word-of-mouth. In the texts of existing research on tourism destination's IWOM, online word-of-mouth interesting and vivid impact on the quality of word of mouth are great, illustrating IWOM which has high quality in the customer in the heart of higher quality [12].

Comprehensive quality and quantity of word-of-mouth's related research, Grant [9], proved that the total effect of word-of-mouth quantity and quality is the most significant, especially the quality of word of mouth, which is different from the path on the purchase intention. So the theoretical model of its influence on purchase intention action is (as shown in Fig. 2).

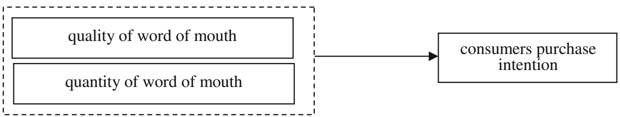


Fig. 2. Word-of-mouth and purchase intention model

This basis the paper puts forward the following two assumptions:

Hypothesis 1: Consumers purchase intention is more significantly affected by the quality of word-of-mouth in far future and the quantity of word-of-mouth influence nearly future purchase intention of consumers more prominently.

Hypothesis 2: When there is far social distance between receivers and senders, the receivers' purchase intention is more significantly affected by the quality of word-of-mouth; When they have nearly social distance, customers' purchase intention is more significantly influenced by the quantity word-of-mouth.

3.2 Design of Experiment

(1) Subjects

The participants were senior undergraduate and graduate students, between the age of 21 to 25 years old, a total of 110 people, is divided into two groups, each group divided into 55 people. There are 28 women, and 27 men in each group. Choosing college students as subjects is mainly based on the following reasons: first, according to the survey of 2013, we found that 85.43% of college students surveyed said "I like traveling". And college students often use the Internet and like to travel. Second, college students and postgraduates with high homogeneity, can effectively control the contrast group and control group of demographic heterogeneity.

(2) Experiment Preparation

Experiment on all testers to get on the Internet first hand tourist information platform and channel was tested. These platforms for tourist destination's IWOM, including Ctrip, Qyer and Ma Feng Wo, the tourism BBS, post tourist experience. There are 90.9% testers, who collect information from those platform. 89.1% from weibo, 60% form We-chat's moment and subscription account, 1.8% from other ways. Due to the W-chat moment's privacy, word-of-mouth authors have close social distance with word-of-mouth readers. So that the experience of travel sites post and micro-blog platform simulation scene, which measures the participants' purchase intention in the different temporal and social distance.

First of all, by reading the study of word-of-mouth, the author summarized main export tablets in quality and quantity of several attributes. Word-of-mouth quantity evaluation is mainly from three aspects [18]: ① multiple online platform to see multiple destinations related IWOM appear; ② the common platform to see multiple destinations IWOM; ③ the IWOM of destination has a lot of forwarding. Measurement of word-of-mouth quality is from the relevance, integrity, interesting and professional comprehensive evaluation.

On this basis, the author collected the related tourism destination online network of word of mouth. To avoid the destination names, which interfere the subjects of purchase intention, we use A and B represent to tourist destination, and both are simulating the ancient town tour relevant scenic spots, this ruling out consumer preferences' disturbance. The IWOM of Destination A reflect word-of-mouth quantity, and tourist destination B's IWOM reflect word-of-mouth quality. Word of mouth simulates the situation in three network platforms like weibo, ctrip and Ma Feng Wo. The tester would be shown 12 items IWOM about A and 3 items IWOM about B. The IWOM about B is 1 item IWOM in each platform, and the IWOM about A is 4 items IWOM in each platform. At the same time the experiment obscures the names of related sites.

(3) Experimental Design

Experiment is divided into two steps: the first measuring different temporal distance and the second experiment measuring different social distance. Measuring subjects for tourism destination A and B different purchase intention, and the scale basis by using three items, designed by Mr Schiff (Schiffman et al.) [17]: ① I am willing to travel to the area; ② I would like to recommend to others the region tourism; ③ I'd be happy to try to travel to the area.

Before answering questionnaire, subjects need to read the IWOM which simulates the IWOM's model of tourism BBS and Micro-blog. Scale takes 7 likert scale from "strongly disagree" to "strongly agree" assignment 1 to 7, recycling questionnaire by examining Cronbach's Alpha $= 0.767 > 0.7$, certificating validity within a reasonable range.

Because tourism is different from the retail purchase, it needs to cooperate with holiday, so the writer has been testing the latest holiday time for the recent test used in measuring temporal distance experiment. One is "weekend" and another is "one year", which are two different time distance measurements. Experiment announced by two word of mouth the identity of senders, 53-years-old workers and 23-year-old college student, respectively.

3.3 The Results

Experiment 1: using SPSS19.0 for data analysis and processing. In the study of the demographic variables such as gender, age did not produce any significant effect, therefore not included in the model. Mixed repeat test analysis of variance, the BOX 'sM inspection results $F(2, 108) = 0.174 > 0.05$, co-variance matrix through the homogeneity test, in the error variance equality Levene test, word-of-mouth quality $P = 0.937$, word-of-mouth quantity $P = 0.094$ were greater than 0.05 through inspection. Through the following Table 1 descriptive statistics analysis, it can be seen that word-of-mouth obvious interaction between quality and quantity exists.

Internet word of mouth and the interaction of the distance and time ($Eta = 0.645$, $p > 0.645$) can better account for the variation model. As shown in Table 1 and Fig. 3, in the distant future buying situation, the purchase intention of subjects who read the IWOM of tourism destination B (reflect word-of-mouth quality), are significantly higher than that of tourism destination A (reflect word-of-mouth quantity); In near future purchase situations, the purchase intention of testee who read destination A (reflect word-of-mouth quantity), is significantly higher than of destination B (reflect word-of-mouth quality). *Hypothesis 1* is validated.

Because of social distance perception differences, the experiment by controlling the reviewers with different background information realizes the control of

Table 1. The results of experiment one

Temporal distance	Online-word-of-mouth				Paired-sample t test	
	Quality		Quantity		t-value	Degree of freedom
	Mean	Standard deviation	Mean	Standard deviation		
Near future	3.18	1.335	4.35	1.377	3.734***	54
Distant future	4.24	1.319	2.47	1.052	7.251***	54

*** is $p < 0.001$

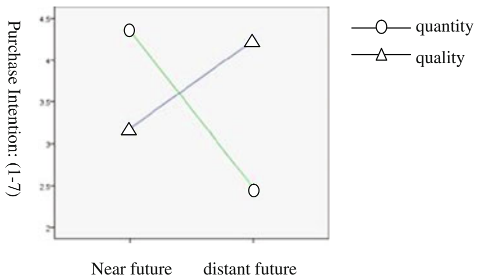


Fig. 3. Impact of word-of-mouth quality and quantity on purchase intention of consumers in different temporal distances

social distance. In the social distance situation in the group, we told the participating commentators, 53-year-old workers (social identities and differences in subjects); Then in the social situation in the group closer, we told participating reviewers, 23-year-old college students (high degree of social identity is similar to the participants).

Social distance measure takes the following scale: (1) I think of myself and word-of-mouth authors similar; (2) I think of myself and word-of-mouth authors psychological distance; (3) I think of myself and word-of-mouth authors belonging to the same group. Adopting likert scale at 7 level produces The result of social distance is that a 23 of college students ($M = 5.26$) is significantly higher than a 53 office worker ($M = 3.44$), showing that the social distance control experiment is effective.

BOX'sM inspection results $F(2, 108) = 0.100 > 0.05$ in Levene test, word-of-mouth quality $P = 0.803$, the word-of-mouth quantity ($P = 0.07$ were greater than 0.05 through inspection. And the interaction of the Internet word of mouth and social distance ($Eta = 0.531$, $p > 0.531$) can better explain the variation model.

As shown in Table 2 and Fig. 4, when the social distance is far between the senders and receivers of IWOM, the subjects' purchase intention of tourism destination B (reflect word-of-mouth quality), are significantly higher than that of tourism destination A (reflect word-of-mouth quantity); When the reader feel

Table 2. The results of experiment two

Social distance	Online-word-of-mouth				Paired-sample t test	
	Quality		Quantity		t-value	Degree of freedom
	Mean	Standard deviation	Mean	Standard deviation		
Far social distance	3.55	1.212	4.51	1.611	5.185^{***}	54
Close social distance	4.80	1.311	2.65	1.181	6.885^{***}	54

*** is $p < 0.001$

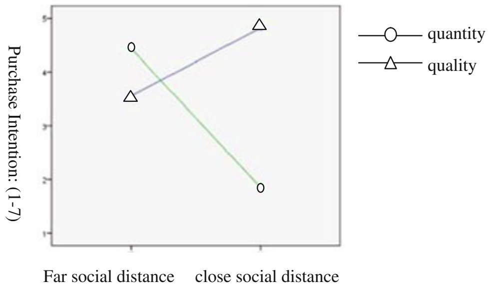


Fig. 4. Impact of word of mouth quality and quantity on purchase intention of consumers in different social distances

they have close social distance, which means that they believe that they have a similar social identity, the purchase intention of destination A (reflect word-of-mouth quantity), is significantly higher than B ((reflect word-of-mouth quality). *Hypothesis 2* is validated.

4 Results, Discussion and Limitations

In this paper, tourism destination online word-of-mouth can be divided into two different measures, including the quality and the quantity of word-of-mouth. This paper, through two experiments, tests the quantity and quality of IWOM influencing on consumer purchase intention under different temporal distance and the social distance situation. The conclusion is as follows:

(1) Quality of Internet word-of-mouth for the forward travel consumer buying may exercise a greater effect and the quantity of IWOM influence on recent purchase intention of consumers more prominently. This suggests that if the time interval between reading and buying is far, the customers' decision-making depends on IWOM quality. It because the customer's high level interpretation will be easily activated, when the cognitive and purchase time interval is far. That means high levels of explanation of IWOM more easily

by cognitive quality and low levels of explaining IWOM quantity will be weakened.

(2) The far social distance between the word-of-mouth authors and readers, consumers purchase intention is influenced significantly by the quantity of word-of-mouth. This suggests that when customers and word-of-mouth authors are far in social distance, the customers' decision-making depends on IWOM quality. It because the customer's high level explanation is easy to be activated. That means high levels of explanation of IWOM quality theory is more easily cognitive and low explaining IWOM quantity will be weakened.

(3) The average of quality of word of mouth influencing on consumers' willingness to buy is higher than word quantity, which means the high-quality IWOM have more extensive and more lasting influence.

In this paper, the study of online word-of-mouth, expands the research contents of psychological explanation level about social distance, and also provides inspiration for tourist destination choice.

First, the relevant tourism destination management on tourism propaganda should choose who have closer social distance with destination target customers. At the same time, the high quality tourism strategy and the relevant evaluation can attract more consumers. The platform of their own tourism official weibo, also should share more high-quality tourist guides, and, release illustrated weibo.

Second, when it is tourist season, the relevant tourism management department can propaganda through multiple platforms, and form a certain scale of word-of-mouth, to attract more attention from people.

Finally, because the study take the simulation experiment method, in the simulation of situation and realistic scenarios exist certain differences. In the future, we should be in the real scene to verify this conclusion.

Acknowledgement. This paper is supported by the Shaanxi Province Social Science Fund (No. 2016R015) and Research on Promoting to Construct Cultural Powerhousee by Cultural-oriented Tourism–Based on Customer Marketing (No. 2016R015).

References

1. Arndt J (1967) Role of product-related conversations in the diffusion of a new product. J Mark Res 4(3):291–295
2. Baber A, Ramayah T et al (2015) Online word-of-mouth antecedents, attitude and intention-to-purchase electronic products in Pakistan. In: USENIX conference on USENIX technical conference, pp 307–318
3. Bailey AA (2004) Thiscompanysucks.com: the use of the internet in negative consumer-to-consumer articulations. J Mark Commun 10(3):169–182
4. Balaji MS, Khong KW, Chong AYL (2016) Determinants of negative word-of-mouth communication using social networking sites. Inf Manage 53(4):528–540
5. Bar-Anan Y, Liberman N, Trope Y (2006) The association between psychological distance and construal level: evidence from an implicit association test. J Exper Psychol Gener 135(4):609
6. Bussière D (2015) Evidence and implications of electronic word of mouth. Springer

7. Chu SC, Kim Y (2011) Study on the relationship between the feature of microblog's word-of-mouse and consumer's brand attitude. Int J Advertising 30(1):47–75
8. Dong D (2012) Contrast analysis of word-of-mouth, internet word-of-mouth and word-of-mouse. Chin J Manage 3:428–436 (in Chinese)
9. Grant R, Clarke RJ, Kyriazis E (2007) A review of factors affecting online consumer search behaviour from an information value perspective. J Mark Manage 23(5):519–533
10. Guo G (2007) The impact of word-of-mouth on consumer's attitude: a theoretical model. Manage Rev 6:20–26 (in Chinese)
11. Haixia Y (2015) The distribution of word-of-mouth across websites and online sales: sensitivity analysis on entropy and network opinion leaders based on artificial neural network. Econ Manage J 10:86–95 (in Chinese)
12. Haiyan C (2011) Study of communication on electronic word-of-mouth of destination. Master's thesis, Wuhan University (in Chinese)
13. Hennig Thurau T, Gwinner KP et al (2004) Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet. J Interact Mark 18(1):38–52
14. Jinjing X (2015) The role of temporal distance and social distance in the consumer decision-making: evidence from eye movements. Master's thesis, Jinan University (in Chinese)
15. Kah JA, Choongki L, Seonghoon L (2016) Spatial-temporal distances in travel intention-behavior. Ann Tourism Res 57:160–175
16. Liu Y (2013) Word of mouth for movies: its dynamics and impact on box office revenue. J Mark 70(3):74–89 (in Chinese)
17. Schiffman LG, Leslie LK (2000) Consumer behavior. Prentice Hall, Upper Saddle River
18. Schubert P, Selz D (1999) Web assessment - measuring the effectiveness of electronic commerce sites going beyond traditional marketing paradigms. In: Hawaii international conference on system sciences, p 5040
19. Trope Y, Liberman N, Wakslak C (2007) Construal levels and psychological distance: effects on representation, prediction, evaluation, and behavior. J Consum Psychol 17(2):83–95
20. Wenming Z, Xu W, Chang F (2014) Relation between electronic word of mouth and purchase intention in social commerce environment: a social capital perspective. Nankai Bus Rev 04:140–150 (in Chinese)
21. Xuan Z (2012) The influence of temporal distance on consumers' online impulse buying behavior. Master's thesis, Huazhong University of Science and Technology (in Chinese)
22. Zainal NTA, Harun A, Lily J (2017) Examining the mediating effect of attitude towards electronic words-of mouth (eWOM) on the relation between the trust in eWOM source and intention to follow eWOM among Malaysian travellers. Asia Pacific Management Review
23. Zheng C, Han Q, Wang H (2015) How do paid posters' comments affect your purchase intention. Nankai Bus Rev 01:89–97 (in Chinese)
24. Zhu X, Li A (2014) Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites. PhD thesis, Guangdong University of Technology (in Chinese)
25. Zou P, Hao L, Li Y et al (2014) Impacts of perceived temporal distance-based sales promotion on purchasing behaviors under information asymmetry. J Manage Sci 01:65–74 (in Chinese)

Analysis and Prediction of Population Aging Trend Based on Population Development Model

Jiancheng Hu(✉)

College of Applied Mathematics, Chengdu University of Information Technology, Chengdu 610065, People's Republic of China
jianchenghu@163.com

Abstract. In recent years, with the acceleration of population aging, the development of population structure is an important factor that influences economy and social development of China. This paper studied the future trend of population in Sichuan of China by establishing a population development equation. Taken the statistical data of census in 2010 in Sichuan, some important factors, such as survival rate, fertility mode and gender ratio etc. are considered to estimate the population structure, especially the proportion of older people, in Sichuan. The model is utilized to predict the population aging trend and aging index in the case of different total fertility rate and to provide reference for the government to make corresponding social and economic decisions.

Keywords: Population aging · Population development equation · Aging index · Total fertility rate

1 Introduction

Population aging is a phenomenon that occurs when the median age of a country or region increases due to rising life expectancy and/or declining fertility rates. Population aging is becoming one of the global problems of the modern times, but with different features from region to region and from country to country. Recent declines in fertility rates and increases in life spans are producing a significant shift in the age distribution of population [4]. Aging of population is closely associated with increased life expectancy, which was mainly determined by better quality health services, new discoveries in pharmaceutics, children diseases treatment. Recent research by the United Nations reveals that the proportion of people aged 60 and over is growing faster than any other age group and it is expected to reach 1 billion by 2020 and almost 2 billion by 2050 (representing 22% of the global population). The proportion of individuals age 80 or over is projected to rise from 1 to 4% of the global population by 2050 [5]. Figure 1 shows the change in the age structure of the world's population over time. Each shaded section reveals the age distribution at one point in time. The elderly share will be much higher in 2050 than it is now.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_27

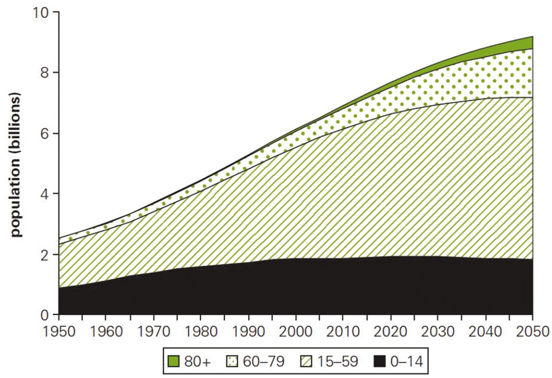


Fig. 1. World population by age group in 1950–2050

Population aging is one of the important factors that influence social and economic development of a country. Since different age groups have different economic needs and productive capacities, a country's economic characteristics may be expected to change as its population ages. Therefor correct forecast on the future population trend has significant guiding to the economic plan of overall country and local government.

In recent years, many articles have been paying more and more attention to population aging prediction and family policy to deal with the aging population problem, such as using gray dynamic model with equal dimension forest theory [9], overlapping generations model [1,7], logistic population model [2], predator-prey population model [8,11,12], stochastic differential equations [10] and so on. In order to analyze and predict the population trend of Sichuan, China, population development equation model is presented and applied to analyze the population of China [3,6].

The United Nations' old standard maintains that a country with more than10% of its population over 60 years old is an aging society. The new standard presents the population of 65 years and above people exceeds 7% of the total population. In fact, according to census of population, there were 88.11 million population aged over 65 years in Chinese mainland, accounting for 6.96% of the total population till November, 2016; while the number turned to 130 million and 10.20% when it came to the population aged over 60 years.

This paper takes the 6th census data in Sichuan as the accordance, and uses a population development equation model to predict the population development in Sichuan. The model is utilized to analyse the population aging trend in the future in Sichuan from a short period, and further predict the long-term population development trend and aging population change condition in Sichuan in the case of different total fertility rate.

2 Population Development Equation

2.1 Continuous Model of Population Development Equation

Let $p(r,t)$ expresses the population density of r years in the year of t, considering the population in the age interval $[r, r+\Delta r)$ from year of to $t+\Delta t$, survival population turns into age interval $[r+\Delta t, r+\Delta r+\Delta t)$, and the number of dead population is $\mu(r,t)p(r,t)\Delta r\Delta t$, which $\mu(r,t)$ is mortality rate of r years population in the year of t. Then we have

$$p(r,t)\Delta r - p(r+\Delta t, t+\Delta t)\Delta r = \mu(r,t)p(r,t)\Delta r\Delta t \tag{1}$$

or

$$\frac{p(r+\Delta t, t+\Delta t) - p(r, t+\Delta t)}{\Delta t} + \frac{p(r, t+\Delta t) - p(r,t)}{\Delta t} = -\mu(r,t)p(r,t), \tag{2}$$

when time increases Δt, the age also increase Δt, that is $\Delta t = \Delta r$. Let $\Delta t = \Delta r \to 0$ in Eq. (2), then

$$\frac{\partial p(r,t)}{\partial r} + \frac{\partial p(r,t)}{\partial t} = -\mu(r,t)p(r,t). \tag{3}$$

Given the initial and boundary conditions, that is, initial population density function $p(r,0) = p_0(r)$ and fertility rate $p(0,t) = f(t)$, the continuous model of population development equation can be obtained

$$\begin{cases} \frac{\partial p(r,t)}{\partial r} + \frac{\partial p(r,t)}{\partial t} = -\mu(r,t)p(r,t), & 0 \le r \le r_m, t \ge 0 \\ p(r,0) = p_0(r), & 0 \le r \le r_m \\ p(0,t) = f(t), & t \ge 0 \\ p(r_m,t) = 0, & t \ge 0, \end{cases} \tag{4}$$

where the r_m is the maximum age of population. Equation (4) is difficult to solve, but in ideal situation, such as stable society circumstance, short time interval etc., the mortality rate and fertility rate are independent with time t, the solution of Eq. (4) can be obtained.

2.2 Discrete Model of Populat1ion Development Equation

In order to study the dynamical population of different ages at any time, the Eq. (4) should be discreted respect to variables r and t. Based on Eq. (1), the r years population $N_r(t)$ in the year of t can be written

$$N_{r+1}(t+1) = s_r(t)N_r(t), r = 0, 1, 2, \cdots r_m - 1 \tag{5}$$

and by virtual of initial condition, also have

$$N_0(t) = s_0(t)b(t), \tag{6}$$

where $s_r(t)$ and $s_0(t)$ are the survival rate of r years population in the year of t and infant survival rate respectively, which can be calculated by the statistical data of population census. $b(t)$ is the population born in the year of t. Concerning the reproductive age interval $[r_1, r_2]$, $b(t)$ can be described as

$$b(t) = \sum_{r=r_1}^{r_2} b_r(t)k_r(t)N_r(t), \tag{7}$$

where $b_r(t)$ is the fertility rate of childbearing age female at r years old in the year of t, i.e. the average number of infant born by each r years old female in the year of t, $k_r(t)$ is gender ratio, the ratio of r years old female in the total population in the year of t.

By representing the population of each age in the year of t as a vector of size r_m, let

$$N(t) = [N_0(t), N_1(t), \cdots, N_{r_m}(t)]^T, \tag{8}$$

the population development Eqs. (5) and (6) can be written as

$$N(t+1) = A(t)N(t) + \beta(t)B(t)N(t), \tag{9}$$

where

$$A(t) = \begin{bmatrix} 0 & 0 & \cdots 0 & 0 \\ s_1(t) & 0 & \cdots 0 & 0 \\ 0 & s_2(t) & \cdots 0 & 0 \\ \vdots & \vdots & \ddots \vdots & \vdots \\ 0 & 0 & 0 \quad s_{r_m-1}(t) & 0 \end{bmatrix}, \tag{10}$$

$$B(t) = \begin{bmatrix} 0 \cdots 0 & b^*_{r_1}(t) & \cdots & b^*_{r_2}(t) & 0 \cdots 0 \\ 0 \cdots 0 & 0 & \cdots & 0 & 0 \cdots 0 \\ \vdots \ddots \vdots & \vdots & \ddots & \vdots & \vdots \ddots \vdots \\ 0 \cdots 0 & 0 & \cdots & 0 & 0 \cdots 0 \end{bmatrix}, \tag{11}$$

$$b^*_r(t) = s_0(t)s_r(t)k_r(t)h_r(t), r = r_1, \cdots, r_2, \tag{12}$$

and the total fertility rate

$$\beta(t) = \sum_{r=r_1}^{r_2} b_r(t), \tag{13}$$

which is the average number of infants born by each female in the year of t. Define the female's fertility mode $h_r(t)$, which is r years old female's fertility weighted factor and satisfied $\sum_{r=r_1}^{r_2} h_r(t) = 1$. Then $b_r(t)$ can also written as

$$b_r(t) = \beta(t)h_r(t). \tag{14}$$

3 Model Establishment and Solution

As Sichuan has come into aging society, in order to make clearly the future population aging trend, we use population development equation to predict the population in a short period and analyze the change trend of future population aging speed. At the same time, predicting the long-term population development process of Sichuan in the case of different total fertility rate, and trying to seek for a relatively good total fertility rate not only can control too fast population growth but also ease the population aging condition.

3.1 Prediction of Population Structure

As the data information provided by the 6th census in Sichuan is limited, and the factors that influence population development are various, it is relatively difficult to get the rule that mortality rate function and fertility mode changes as the time changes. The population fertility policy will not change in a relatively short period; in the condition of stable society, we usually suppose that survival rate $s_r(t)$, fertility mode $h_r(t)$ and gender ratio $k_r(t)$ relatively do not change. The data published from the 6th census in Sichuan shows the total fertility rate of women $\beta(t) = 1.075$ at present. Under the supposition above, we use the data of the 6th census in 2010 in Sichuan as the cardinal number to predict the population in the future six years in Sichuan. The prediction result is shown as Table 1.

From the prediction results, we can see that the population in Sichuan is still increasing. After Guangdong, Shandong and Henan, Sichuan is ranked fourth big province in the whole country population, also aging serious provinces. Although the data of the census shows the population fertility rate is relatively low, and the total population still continuously grows.

From Table 2, we can see from the calculation result of population proportion at each age section that the proportions of aged people and children still continuously increase, but the proportion of adults continuously decreases. In 2010, people under the age of 14 years accounts for the total population of 16.97%, 0.37% points higher than the national average level; 15–64 year old population account for the total population of 72.08%, 2.45% points lower than the national average; the population aged 65 years account for the province's total population of 10.95%, 2.08% points higher than the national average.

3.2 Aging Index

The aging index (AI, sometimes referred to as the elder-child ratio) is a common indicator of changes in age structure, which is defined as the number of people aged 65 and over per 100 young people aged 0–14 years [10]. In 2000, only a few countries (Germany, Greece, Italy, Bulgaria, and Japan) had more elderly than youth (aging index above 100). To date, aging indexes are much lower in developing countries than in the developed world, but the proportional rise in the

Table 1. Age structure of population by five-year age group, 2010–2016

Age structure (years)	2010	2011	2012	2013	2014	2015	2016
0–4	4277055	4053597	3948212	4059425	4214668	4781547	5427837
5–9	4277626	4293234	4359772	4421205	4389209	4264325	4041644
10–14	5092442	4832421	4597470	4347174	4310136	4268574	4284142
15–19	6213182	6240489	6147844	5893439	5505725	5080985	4821552
20–24	6360027	6240153	6047303	6201402	6234291	6193317	6220385
25–29	4514128	4779065	5055472	5395750	5886962	6333555	6214086
30–34	4874122	4529418	4569747	4499193	4389343	4489706	4753438
35–39	8077512	7665262	6998152	6248575	5514606	4839660	4497725
40–44	7819390	7982458	8320171	8225243	8268743	7997893	7589384
45–49	5934098	6988192	7385771	7423684	7470576	7707164	7868020
50–54	4142765	3507478	3346675	4030751	4908537	5809492	6838453
55–59	5725272	5805843	5583171	5302779	4637600	4015098	3398585
60–64	4304402	4414855	4770329	5089829	5300649	5467686	5541134
65–69	3295231	3405761	3555960	3645737	3813371	4004605	4106617
70–74	2408632	2489142	2581438	2698645	2870848	2914706	3013052
75–79	1588550	1682718	1733306	1774822	1836206	1960249	2027791
80–84	962223	1006531	1054057	1100412	1124821	1123230	1192039
85–89	399472	423506.4	456082.5	502784.8	519828.3	557905.7	582064.4
90–94	118758	128692.8	137625.8	146182.7	161883.4	172212.2	182358.7
95–99	29414	31314.39	32530.13	34764.69	36467.73	40232.09	43599.68
100+	3227	1706.79	1944.118	1895.151	2065	2130.435	2332.162
Total	80417528	80501838	80683033	81043692	81396535	82024275	82646242

Source: Authors' calculations based on data in the 6th population census 2010.

Table 2. Proportion of the population at each age section in total population

Age structure (years)	2010	2011	2012	2013	2014	2015	2016
0–14	13647123	13179252	12905455	12827804	12914013	13314446	13753624
(%)	−16.97%	−16.37%	−16.00%	−15.83%	−15.87%	−16.23%	−16.64%
15–64	57964898	58153213	58224635	58310644	58117031	57934558	57742763
(%)	−72.08%	−72.24%	−72.16%	−71.95%	−71.40%	−70.63%	−69.87%
65+	8805507	9169373	9552943	9905244	10365491	10775272	11149855
(%)	−10.95%	−11.39%	−11.84%	−12.22%	−12.73%	−13.14%	−13.49%

Source: Authors' calculations based on data in the 6th population census 2010

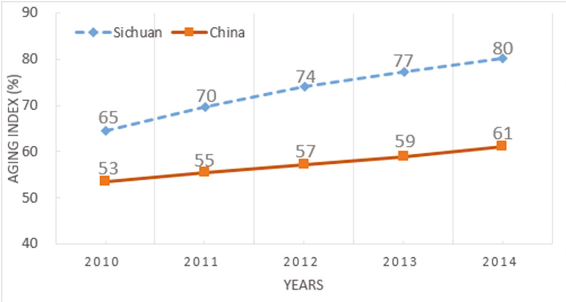


Fig. 2. Aging index of Sichuan and China in 2010–2014

aging index in developing countries is expected to be greater than in developed countries.

In 2010, the AI in China was approximately 53, meaning that there were 53 people aged 65 and over for every 100 persons aged 0–14 years, while the AI in Sichuan was approximately more than another 12. As Fig. 2 shown, the AI in both China and Sichuan increased progressively from 2010 to 2014, and the increment of AI in Sichuan is more than that of China in studied period from 65(2011) to 80(2014).

3.3 Total Fertility Rate Analysis

In the late-1970s started the family planning policy which implemented, especially in the urban only child policy, China effectively was suppressing the population growth, the fertility rate dropped continually, and in recent years, the fertility rate has remained at 1.4%–1.5% level for a long time, which have also accelerated the aging of population. In order to promote the balanced development of population, adhere to the basic state policy of family planning, improve the population development strategy, China fully implement two-children policy, that is, a couple can have two children. Two-children policy appropriate to slow down the population aging process, help to solve the aging population.

From the short term, the family planning policy will not influence the population structure of Sichuan temporarily, however, from the long term, it will have played a positive role on easing the aging population. In the population analysis and decision-making, the total fertility rate is an important index and parameter for the prediction of population. It can be directly used to compare different periods of women's fertility level, which has important reference to the long-term macro decision of the population. In order to study population fertility policy, the following supposes that the other factors do not change in the future years, and predicts the population structure of Sichuan in different total fertility rate. Following, we will suppose that the total fertility rate $\beta(t)$ is respectively at 1.5, 2.0 and 2.5, and predict the total population and the proportion of 65+ years old population in the future 50 years in Sichuan. The prediction results are shown as Figs. 3 and 4.

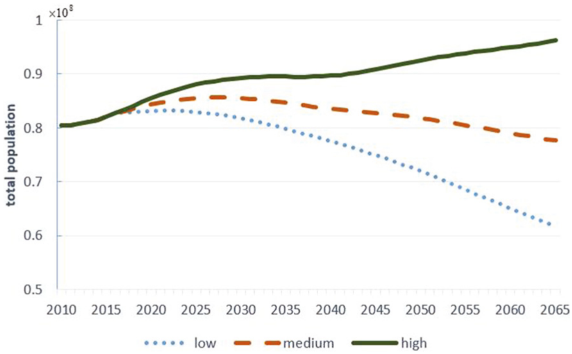


Fig. 3. Prediction of total population in 2010–2065

From Fig. 3 we can find that the total population quantity in the future will increase continuously in the case of high total fertility rate. When the total fertility rate maintains the low level, the total population has a short term growth, and soon will decrease steadily. In the case of medium total fertility rate, the total population will maintain stable in the future 50 years. Meanwhile, the prediction of aging population proportion in the future 65 years is shown as Fig. 4.

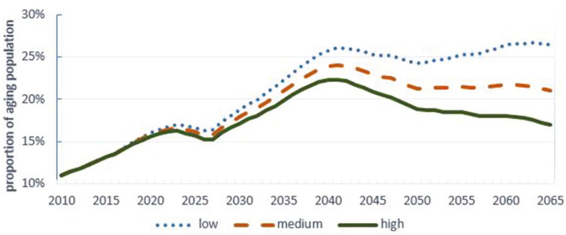


Fig. 4. Proportion of population at 65+ years old in 2010–2065

As Fig. 4 shown, in the case of low total fertility rate, the proportion of aging population will continuously increase and finally exceed 25% in 2060. In the case of medium total fertility rate, the aging population increase slowly, and stabilizes around 22% after 30 years. When the total fertility rate is high, the proportion of aging population is relatively slow, and gradually stabilize and maintain in the below 20% after 30 years of growth.

From the prediction results, we can find that the relatively high fertility rate will make the population aging keep at a relatively low level, but will cause the rapid population growth. For example, considering the proportion of aging population and the total population in the period of 2020–2040, the proportion of aging population changes from 16% to 26% and the total population decreases from 83 million to 77 million in the case of low total fertility rate. Meanwhile, the proportion of aging population only increases 6% from 16% and the total population increases 13 million from 83 million in the case of high total fertility rate. The prediction indicates that only when the total fertility rate keeps at 2.0 can a relatively good result appear.

The rate of population aging may also be modulated by migration. Immigration usually slows down population aging, because immigrants tend to be younger and have more children. On the other hand, emigration of working-age adults accelerates population aging. The 6th census shows that in 2010 the population mobility rate of Sichuan was 18.92%, lower than the national average only 0.59% points. Among them, the emigration rate of Sichuan was 23.24%, and the immigration rate was 14.59%.

4 Conclusions

This paper predicts the population structure in the future six years in Sichuan by establishing the population development equation. And the results show that total population in Sichuan will continuously increase and the aging population is increasingly aggravated. The population aging is an inevitable population problem. The aging population increase will be a series of problems to society, we can not only solve it through population policy but also old-age supporting policy, and social security policy, and regulating economic structure, etc.

In order to analyze the effect of fertility rate on the aging of the population, we predict the population structure in the future 50 years in Sichuan in the case of different fertility rate. And the results indicate that appropriate fertility rate will help to regulate population structure and ease the population aging. When the total fertility rate is 2.0, the population development in the future in Sichuan will be relatively stable and the trend of population aging will slow down and finally keep at the level of 22%.

Acknowledgements. This research is supported by the Scientific Research Foundation of CUIT (Grant No. KYTZ201425) and Center of System science and Enterprise Development Research (Grant No. Xq16C01).

References

1. Abdessalem T, Cherni HC (2016) Macroeconomic effects of pension reforms in the context of aging populations: overlapping generations model simulations for tunisia. Middle East Dev J 8(1):84–108
2. Alves CO, Delgado M et al (2015) Existence of positive solution of a nonlocal logistic population model. Zeitschrift fr angewandte Mathematik und Physik 66(3):943–953
3. Cui Q, Lawson GJ, Gao D (1984) Study of models of single populations: development of equations used in microorganism cultures. Biotechnol Bioeng 26(7):682–686
4. DESA (2013) World population aging 2013. Technical report, PD Department of Economic and Social Affairs (DESA), New York

5. DESA (2013) World population prospects: the 2012 revision, key findings and advance tables. Technical report, PD Department of Economic and Social Affairs (DESA), New York
6. Liu M, Liao D et al (2013) The correctional model of population development equation. Model Numer Simul Mater Sci 03(4):139–141
7. Muto I, Oda T, Sudo N (2016) Macroeconomic impact of population aging in Japan: a perspective from an overlapping generations model. IMF Econ Rev 64:408–442
8. Novozhilov AS (2004) Analysis of a generalized population predator-prey model with a parameter of normal distribution over individuals of the predator population. J Comput Syst Int 43(3):378–382
9. Shi LN, Huang HM et al (2009) Aging population trend prediction based on the gray recurrence dynamic model with equal dimension. Ludong Univ J 25(4): 315–317
10. Yang F, Zhang LD, Shen JF (2013) Further discussion on population growth models by stochastic differential equations. Adv Mater Res 705:499–503
11. Zhang Y, Gao S et al (2015) On the dynamics of a stochastic ratio-dependent predator-prey model with a specific functional response. J Appl Math Comput 48(1–2):441–460
12. Zhou A, Sattayatham P, Jiao J (2016) Analysis of a predator-prey model with impulsive diffusion and releasing on predator population. Adv Differ Eqn 1:1–18

Evaluation of Progressive Team Intervention on Promoting Physical Exercise Behavior

Xinyan Guo(✉)

Economy and Management Faculty in Chengdu Sport Institute,
Chengdu 610041, People's Republic of China
sindystory@163.com

Abstract. Focusing on parts of urban residents, which were divided into two groups, intervention group (165 people) and control group (178 people), a "5As" physical exercise behavior intervention process (Ask, Access, Advice, Assist and Arrange) was applied to the intervention group people with auxiliary materials including intervention manuals and exercise prescriptions. The baseline data and the tracking data of the two groups were collected and analyzed by descriptive statistics, comparison of means, analysis of variance and correlation analysis to study the effectiveness of intervention. The results of tracking survey showed that residents had better acceptance for the use of intervention. Before and after the intervention, there was a certain change about the residents' physical exercise behavior and there were significant differences in the behavior variables ($P < 0.05$). Physical exercise behavior belongs to preventive innovation which requires corresponding health education and intervention. The research adopts a step by step mode to implement targeted intervention measures, and peer groups make it easier to accept and prove to be efficient for the individuals. And the research provides a new idea and method for the physical exercise intervention.

Keywords: Physical exercise behavior · Progressive intervention · Tracking survey · Urban resident

1 Introduction

In the modern society, sports and human survival and development are more and more closely related, its significance and function far exceeded the original concepts and categories. Physical activity can effectively change people's way of life, prevent all kinds of chronic diseases, and moderate physical activity is one of the most simple ways to improve and maintain health [3,17,18]. People pay much more attention to living environment, living styles and the consequent changes of morbidity and mortality rates [16]. Researches on health management and promotion show a close relationship between residents' living style and physical health.

In Japan, less than 30% of Japanese do frequent (usually one year or more) and regular (30 min per time, twice a week) physical exercises. A survey in

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_28

Australia showed that 53% of the adults didn't take regular exercises with their sitting time being 4.7 h per day or more [10]. High income countries like U.S. and U.K. showed the same results. According to the survey, in U.K., only 39% male and 29% female adults were doing exercises regularly, moreover, the participation rate declined with the change of age, work and family [1,11,12].

There are many exercise behavioral intervention studies. Geoff et al. (2010) conducted a pre-post survey research to 200 British female undergraduates who did not attended any physical exercises, and analyzed the benefits and cognitive disorders [13]. This study indicated that interventions should give priority to solve the disorders of participating physical exercises which can be perceived by individual. On the other hand, interventions should emphasize this cognition that physical exercises are beneficial to health, thereby promoting modifications of cognition and behavior for participating physical exercises. In order to study the influence of PE class to students' physical activities, Sallis et al. designed a trial that it made 338 senior undergraduates to attend randomly intervention credit courses or control credit courses [19]. The results demonstrated that the total volume of activities including leisure time, strength exercises and flexible exercises in female group was increased due to interventions, which facilitated female to participate healthy physical exercises. However, there was not significant in male group, because male students were more like sports than female found in basic survey. About the applied studied of stage-based intervention, Marcus et al. performed an exercise intervention the Imagine Action campaign for community volunteers to research the results between pre-exercises and post-exercises, and this study indicated that 30% of participants in Contemplation and 60% of participants in Preparation moved into Action, while 31% in Contemplation moved into Preparation. This study also indicated that less expensive and intensive exercise intervention can be adopted significantly [14]. Bogdan et al. evidenced the effectiveness of stage-based intervention in their study which aimed at exercises behavioral intervention for 560 orthotic patients [2]. Other researchers also proved the advantages of stages-matched intervention to stages-non-matched intervention using other approaches such as Internet aided intervention [4].

However, to make it function well, coordination and cooperation are needed between individuals, family, society and various departments (such as health, education etc.). At the same time, it also depends on the scientific method of promoting behavior to intervene [5,7].

In this paper, Progressive Intervention is adopted to understand the individual characteristics, strengthen the awareness of individual self-management and help individuals form a good behavior habits. Then through baseline data and tracking data, the effectiveness of intervention was analyzed to provide operational framework and theoretical guidance for the health promotion, especially for the "sub health" status of the population.

2 Objects and Methods

2.1 Participants

Stratified random sampling method was used to select the urban residents participants from the fitness clubs, parks, squares, community focus points etc. And they include volunteer solicitation, recruitment network, and random samples through the local network. Sample size for the baseline survey was 343. Samples were randomly divided into two groups: intervention group and control group, 165 in the intervention group, and 178 in the control group (Tables 1 and 2).

Table 1. Gender of respondents (baseline data) $N = 343$

Gender	%
Male	47.1
Female	52.9

Table 2. Age of respondents (baseline data) $N = 343$

Age	≤18	18–29	30–39	40–49	50–59	≥60
(%)	6.5	10.8	25.7	28.1	19.2	9.7

To ensure the quality of the intervention, staff training was carried out before the test to ensure that the intervention process was completely. And all these works are taken out accordance with the framework of the theory hypothesis.

2.2 Test Instruments

The related testing instruments used in the study include the test scale, the intervention manual and the experiment site, and so on.

The test scale is mainly used to measure the process and the result of the residents' physical exercise behavior. This study integrated the health belief theory and the theory of planned behavior and propose that the exercise behavior outcome of urban resident can be predicted by some factors including behavioral attitude, body norm, perceived subjective control, perceived susceptibility, perceived severity, perceived behavior benefit, perceived behavior disorder, behavior clue and behavior intention. The scale is based on the integration of relevant theory [8]. A self-designed questionnaire was formed on the combination of the theory of planned behavior and health belief model, the questionnaire contain 4 items of personal basic information (including gender, age, occupation, education etc.), perceived susceptibility, perceived severity, perceived barriers, perceived benefits, behavioral cues, behavioral attitude, subjective norm, perceived behavioral control, behavioral intention and behavior, a total of 33 items.

Scale by Li Kete five point scale scoring methods: among them, perceived barriers project assigned the remaining items reverse, positive score, namely "strongly disagree" for 1 points, "disagree" for 2 points, "not sure" for 3 points, "agreed" for 4 points, "strongly agree" for 5 points.

The effectiveness of the physical exercise behavior at the present stage (including the action stage and the intensity of action) and the related characteristics were measured [6].

Based on exercises behaviors traits and stages of residents, the invention handbook was composed of introduction, interpretation and suggestion. The content is as follow:

(1) Health, sub-health and diseases
The aim of this part is to provide some essential concepts about their physical fitness for residents so that lay a good foundation for healthy life style. For example, health can be divided into three states, namely health, sub-health and diseases. A state of sub-health is either non-health or non-diseases. The symptoms of sub-health include lack of energy and other weakness, but these symptoms cannot be diagnosed by clinical examinations.

(2) Introduction of scientific exercise methods
In the part, the benefits of regular exercises will be introduced, including the improvement of physical fitness, adjustment of psychological factors and so on. And then to introduce what is regular exercises, how to control the amount of exercises, how to warm-up and how to eliminate fatigue quickly.

(3) The common exercise methods aerobic exercise
The most common exercise is aerobic exercise that is to improve the cardio respiratory fitness and physical fitness. Furthermore, methods and skills of aerobic exercise will be introduced in this part.

(4) Other exercise methods
These exercise methods do not need complex techniques, fitness facilities, and extra costs, at the same time these methods are effective, feasible and popular for public. For example, jogging, running, bicycle riding, aerobics and fitness path.

(5) Exercise methods of different population
People should perform suitable exercises based on their physical and psychological conditions such as ages, gender and vocation. The main contents of this part is to provide distinct advices for different population.

2.3 Test Procedure

The study is a quasi-experimental design due to the large sample size and the certain range of exercise behavior intervention. In order to ensure the rationality and integrity of data analysis, the empirical research is divided into two stages: the first stage is baseline survey data analysis, and the second stage is tracking data analysis. The test procedure is shown in Fig. 1.

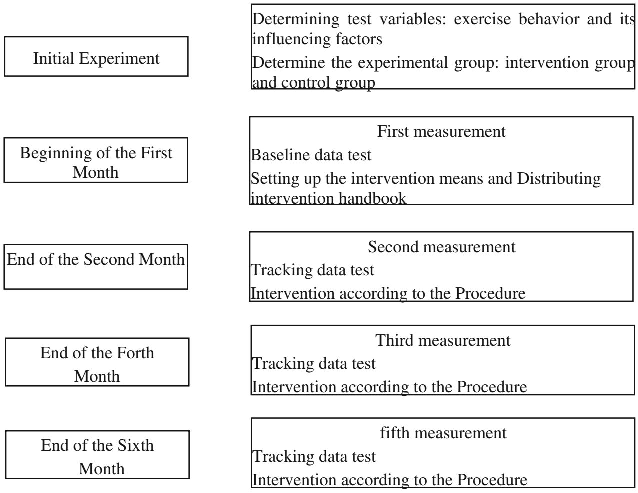


Fig. 1. Intervention procedure

Firstly, the baseline survey was performed with the permission of subject individuals. The questionnaire was distributed to the residents, and the characteristics of the physical exercise behavior and basic information were measured, then the experimental group of residents were carried out by the intervention staff.

After six month behavior intervention, behavioral tests were carried out every two months, four times totally. Both groups use the same questionnaire scale. The survey was conducted in a narrative way of step-by-step introduction, explanation and recommendations.

3 Intervention Process

3.1 Intervention Evaluation

Firstly, Epi Data was used to establish a database, and all the data were checked after the entry. Then the modeling and analysis of the related research can be carried out only after the examination of the normal distribution, in which the reference value of all samples were within the reference range of criteria. And the absolute value of Skewness coefficient and kurtosis coefficient should be in the 0 to 1 range, and the S-W test P value is greater than 0.05, which demonstrated that the formal sample can meet the requirements of normal distribution [15].

Urban residents of Sichuan randomly selected for the survey sample (n = 343), and the two subjects showed no significant difference in the results of exercise behavior and behavioral characteristics variables ($p > 0.05$).

3.2 Follow-up

Complete physical exercise behavior intervention points can be summed up in 5 words: Ask, Access, Advice, Assist and Arrange (5As).

Ask is to inquire about the individual residents to participate in physical exercise and understand their health status; assess is to measure individual residents' intention of participating physical exercise, behavior characteristics and the behavior of the stage; advice is to provide targeted exercise recommendations, support and help based on the individual physical exercise behavior; at last follow-up (shown in Fig. 2).

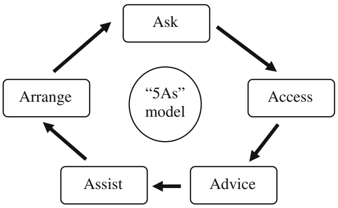


Fig. 2. "5As" Intervention flow chart of physical exercise behavior

After the beginning of intervention, according to the experimental arrangement, "Arrange" is set up in the four follow-up tests. The main purpose is to understand if the residents still persist in physical exercises, congratulate the residents who have maintained physical exercise and encourage them to continue to persist; and give the affirmation to those who maintained a period of time and then give up, and encourage them to start exercise behavior again. At the same time, according to the interview, guidance and help are given, in case of giving up.

4 Intervention Outcome

Based on the theoretical hypothesis, the exercise behavior intervention experiment is implemented to analyze the changes of physical exercise behavior between intervention group and the control group before and after the test period.

Firstly, the slope homogeneity of the covariance analysis is tested. The variance analysis model were analyzed in which the variables of behavioral characteristics after intervention were dependent variables, the group variables (1 = intervention group and 2 = control group) were independent variables. The result showed that the interaction between independent variables and covariates was not significant ($P > 0.05$), that is, the assumption of the slope homogeneity (Table 3).

The result showed that the effectiveness is not significant from previous perceived susceptibility, previous perceived severity, previous perceived benefits, previous perceived barriers, previous behavioral attitude, previous subjective norm and previous perceived behavioral control variable ($P > 0.05$). But some variables' effectivenesses are significant ($P < 0.05$), such as previous behavioral cues and previous behavioral intention. Results are shown in Table 4.

Table 3. Covariance pre-analysis results (interaction)

Source of variance	P value
Group* Previous behavioral cues	0.363
Group* Previous perceived susceptibility	0.627
Group* Previous perceived severity	0.351
Group* Previous perceived benefits	0.094
Group* Previous perceived barriers	0.126
Group* Previous behavioral attitude	0.521
Group* Previous subjective norm	0.438
Group* Previous perceived behavioral control	0.095
Group* Previous behavioral intention	0.766

*The difference is significant at 0.05 level.
**The difference is significant at 0.01 level.

Table 4. Covariance pre-analysis results (Intervention effectiveness)

Source of variance	P value
Previous behavioral cues	0.044*
Previous perceived susceptibility	0.426
Previous perceived severity	0.233
Previous perceived benefits	0.068
Previous perceived barriers	0.091
Previous behavioral attitude	0.084
Previous subjective norm	0.258
Previous perceived behavioral control	0.086
Previous behavioral intention	0.032*

*The difference is significant at 0.05 level.
**The difference is significant at 0.01 level.

Therefore, the exercise behavior variables of the two groups were compared and analyzed (Tables 5 and 6). The results indicated that the exercise behavior of the intervention group was significantly higher than that of the control group, which means the progressive intervention is effective to promote the formation of physical exercise.

Table 5. Comparative analysis of the effectiveness before and after the intervention

Variable	Intervention group	Control group	P value
Perceived susceptibility	$3.35 \pm 0.0.583$	$2.83 \pm 0.0.518$	0.032*
Perceived severity	2.89 ± 0.167	$2.49 \pm 0.0.273$	0.050*
Perceived benefits	$3.05 \pm 0.0.267$	$3.68 \pm 0.0.415$	0.021*
Perceived barriers	$3.58 \pm 0.0.326$	$2.94 \pm 0.0.572$	0.006**
Behavioral attitude	$4.25 \pm 0.0.597$	$3.92 \pm 0.0.486$	0.006**
Subjective norm	$4.15 \pm 0.0.278$	$4.00 \pm 0.0.436$	0.000**
Perceived behavioral control	$4.83 \pm 0.0.352$	$4.70 \pm 0.0.492$	0.000**

*The difference is significant at 0.05 level.
**The difference is significant at 0.01 level.

Table 6. Difference analysis of adjusted mean

Variable	(I) Group	(J) Group	Mean difference (I–J)	P Value
Behavioral cues	Intervention group	Control group	0.42	0.022*
Behavioral intention	Intervention group	Control group	0.36	0.003**

*The difference is significant at 0.05 level.
**The difference is significant at 0.01 level.

5 Discussion

The subject was intervened by progressive intervention, and the effectiveness of the intervention was confirmed by the comparative analysis of the intervention group and the control group.

The influence factors of urban residents' exercising behavior are multidimensional, including 9 factors: behavioral attitude, subjective norm, perceived behavioral control, perceived susceptibility, perceived severity, perceived benefits, perceived barriers, behavioral cues and behavioral intention. Perceived susceptibility and perceived severity of physical exercise have effects on perceived benefits and behavior attitude: perceived susceptibility, perceived severity, behavior barriers and perceived benefits have a positive impact on behavior attitude; perceived benefits have a negative role in behavior barriers but positive in behavior attitude; behavior attitude, subjective norm and perceived behavioral control function positively to behavior intention and behavior attitude does positive role to perceived behavioral control at the same time; behavior intention, perceived behavioral control, perceived benefits, barriers and cues exerts positive effects on behavior itself. Based on the survey results, in the baseline data, the majority of urban residents are aware of the benefits from physical exercise, but they are not necessarily involved in the fitness. In other words, the promoting effect which perceived benefits impacted on behavior intention is still to be strengthened. For there are more worried about, such as ways of physical exercises and strength, physical training environment, time cost, etc.

In addition, although the residents hold a positive attitude for physical exercise benefits, but due to the limited understanding of their own physical condition and the relevant authority on argument, there is no clear subjective feeling for the contrast effect not to participate in fitness activities, and actively participate in fitness activities. But the intervention group was improved in these aspects.

In the four follow-up tests, according to the time schedule, the main purpose set of the "Arrange" is to understand whether the residents are still persist in this behavior or not [9]. In the process of behavior intervention, the people who did not participate in physical exercise should be divided into stages to carry out the work, such as conducting appropriate in-depth interviews with individuals, giving psychological guidance, emphasizing the importance of exercise behavior in their consciousness, at the same time, enhancing the confidence of physical exercise by demonstration role or embodiment contrast, and helping them realize the change of behavior as early as possible. In order to pursue a better cooperation of individual residents, access to information about physical exercise, intervention staffs may need to make some special arrangements for the environment of the conversation. For example, put some health handbooks or posters about physical exercise, display some sports souvenirs on desk, to help residents receive health intervention, and let them feel about that physical exercise is a very important work in the process of communication.

6 Conclusions

This study indicated that individual had different psychological cognitive status and behavior intention level, and faced different difficulties and obstacles. Based on the survey results, in the baseline data, the majority of urban residents are aware of the benefits from physical exercise, but they are not necessarily involved in the fitness. In other words, the promoting effect which perceived benefits impacted on behavior intention is still to be strengthened. For there are more worried about, such as ways of physical exercises and strength, physical training environment, time cost, etc. In addition, although the residents hold a positive attitude for physical exercise benefits, but due to the limited understanding of their own physical condition and the relevant authority on argument, there is no clear subjective feeling for the contrast effect not to participate in fitness activities, and actively participate in fitness activities. But the intervention group was improved in these aspects.

National fitness activities is mass fitness activities, to promote physical and mental health. Urban residents as the main subject of national fitness activities, their views of the community sports, support and participation in decision-making, directly impact on the mass sports development. Using progressive intervention mode can make people to clear the relationship between physical exercise and promote the health, to deepen the understanding of physical exercise behavior. in this study the intervention based physical exercise behavior intervention points (5As), and the result meant progressive intervention is effective to promote the formation of physical exercise.

References

1. Begg S, Khor S, Bright M (2003) Risk factor impact on the burden of disease and injury in Queensland. Queensland Burden of Disease and Injury Circular Series
2. Bogdan RC, Biklen SK (1982) Qualitative Research for Education. Allyn and Bacon, Boston
3. Physical Activity Guidelines for Americans Committee et al (2008) Physical Activity Guidelines for Americans. US Department of Health and Human Services, Washington, DC
4. French KE, Nevett ME (1997) Knowledge representation and problem solution in expert and novice youth baseball players. Res Q Exerc Sport 67(4):386–395
5. Guo X (2015) Application of diffusion of innovation theory on the intervention research of exercise behaviors in urban residents. Mod Prev Med 31(3):202–208
6. Guo X (2015) Construction and Validation of the Integration Model of Planned Behavior Theory and Health Belief Model. Springer, Heidelberg
7. Guo X (2016) Examination on the discontinuity of physical exercise behavior stages and its implication for intervention. J Chengdu Sport Univ 42(3):42–47
8. Guo X, Xu J (2010) Making and testing of the scale to measure the change stage for the exercising behavior of urban residents based on a survey into some Sichuan urban communities. J Chengdu Sport Univ 5:71–74
9. Guo X, Zhang W (2017) Stages and Processes of Self Change of Exercise Behavior: Toward an Integrative Model of Change. Springer, Singapore
10. Harper C, Epidemiologist C (2008) Challenges and opportunities - key challenge 2: make prevention a key to reducing health inequality. In: The Health of Queenslanders 2008-8211; Prevention of Chronic Disease: Second Report of the Chief Health Officer, Queensland. Queensland Burden of Disease and Injury Circular Series
11. Department of Health (2005) Choosing activity: a physical activity action plan. Healthex Specialist
12. Department of Health (2005) Choosing health: making healthy choices easier. Primary Health Care
13. Lovell GP, Ansari WE, Parker JK (2010) Perceived exercise benefits and barriers of non-exercising female university students in the United Kingdom. Int J Environ Res Pub Health 7(3):784–798
14. Marcus BH, Banspach SW (1992) Using the stages of change model to increase the adoption of physical activity among community participants. Am J Health Promot AJHP 6(6):424–429
15. Markus KA (2012) Principles and practice of structural equation modeling by rex b. kline. Struct Equ Model Multidisciplinary J 19(3):509–512
16. Mesbah M (2016) Measurement and analysis of quality of life related to environmental hazards: the methodology illustrated by recent epidemiological studies. Int J Manage Sci Eng Manage 11(2):1–16
17. World Health Organization (2008) Prevalence of insufficient physical activity, age 15+, age-standardized: both sexes. World Health Organization, Geneva
18. World Health Organization (2013) Bull World Health Organ. World Health Organization
19. Sallis JF, Calfas KJ (1999) Evaluation of a university course to promote physical activity: project grad. Res Q Exerc Sport 70(70):1–10

SEM-Based Value Generation Mechanism from Open Government Data in Environment/Weather Sector

Xiaoling Song[1,2](✉), Charles Shen[2], Lin Zhong[3], and Feniosky Peña-Mora[2]

[1] School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094, People's Republic of China
songxiaoling426@163.com

[2] Advanced ConsTruction and InfOrmation techNology (ACTION) Laboratory, Department of Civil Engineering and Engineering Mechanics, Columbia University, New York 10027, USA

[3] Business School, Sichuan University, Chengdu 610065, People's Republic of China

Abstract. Environmental issues are harmful effects of human activities on the biophysical environment. More and more citizens are engaged in the open data movement for different purposes, in particular, emerging private companies are being built on Open Government Data (OGD) to predict extreme weather events for environmental protection, to satisfy more social requirements (e.g. job creation) and to make economic profits, simultaneously. Emerging companies are utilizing OGD to generate values. Building on a synthesis of the OGD literature and established theories of value generation, we develop a structural equation modelling (SEM)-based model to explore the causal relationship between OGD and value generation. This study constructs the conceptual structural equation model and lays a foundation for the upcoming research.

Keywords: SEM · Value generation · Open government data · Environment/weather sector

1 Introduction

Environmental issues are harmful effects of human activities on the biophysical environment. Major current environmental issues may include *climate change*, pollution, environmental degradation, and resource depletion etc. Global warming, or *climate change* as it's now commonly referred to, is the greatest environmental threat we've ever faced. There is little doubt that climate change is contributing to the extreme weather disasters we have been experiencing, such as worsening air quality, searing heat, and more frequent extreme weather events (e.g. raging storms, ferocious fires, severe droughts, and punishing floods). It threatens our health, communities, economy, and national security [7], for example, the climate change-related heat mortality in the first half of 2012 is a part

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_29

of a deadly trend. In 2011, at least 206 people died from extreme heat, up from 138 fatalities in 2010 and nearly double the 10-year average, according to the National Oceanic and Atmospheric Administration [8]. If we don't do more to reduce fossil fuel emissions and other heat-trapping greenhouse gases that are making heat waves more intense, more than 150,000 additional Americans could die by the end of this century due to excessive heat. Heat-related death is just one deadly side effect of extreme weather tied to climate change [7]. Other extreme events also have side effects, such as storms can cause drowning, contaminate drinking water and result in outbreaks of infectious diseases; heat and ozone smog increases respiratory diseases (e.g. asthma) and worsens the health of people suffering from cardiac or pulmonary disease.

According to NRDC (2015), the global warming has mainly six side effects on human health, including more vector-borne disease, more water-borne illness, longer allergy seasons, declining crops, air pollution and more heatwaves [7]. Therefore, as we work to stop climate change and prevent those associated impacts, we are also developing policies to minimize human suffering.

At present, worldwide nations have begun taking steps to combat this growing threat, working toward an international agreement in which every country on earth plays its part. Many of the world's largest polluters have stepped up with significant commitments, amplified by efforts from cities, businesses, sports leagues, non-governmental organizations (NGOs), and many other individuals and groups that have responded to the urgent need for a climate action. For example, since carbon pollution fuels climate change, drives extreme weather, threatens communities and cuts too many lives short, the U.S. Environmental Protection Agency has developed the Clean Power Plan, which sets the first national limits on carbon pollution from power plants and provides states with the flexibility to meet them.

The NRDC has suggested more solutions to address the global warming issues, such as setting limits on global warming pollution; investing in grean jobs and clean energy; driving smarter cars; creating green homes and buildings; as well as building better communities and transportation networks. These solutions will motivate our lawmakers to quit ignoring climate change and start limiting carbon pollution that is heating our planet and increasing the intensity of extreme weather. However, the government cannot be expected to do everything to protect the environment. In cooperating with the private sector as a partner can make sense based on the principle of partnership and appropriate process. Opening data is recognized as a driver of efficiency and a vehicle for increasing transparency, citizen participation and innovation in society. Open data can help improve food security, healthcare, education, cities, the environment, and other public and private services that are critical to development. Taiwan environment protection agency (TEPA) views environmental data as a strategic asset to protect our environment. *At present, more and more citizens are engaged in the open data movement for different purposes, in particular, emerging private* *companies are being built on Open Government Data or Open Environmental Data to predict extreme weather events for environmental protection, to satisfy more social requirements (e.g. job creation) and to make economic profits, simultaneously.*

Today, we are generating and storing more data than at any other time in history. The public sector and especially the various government branches are one of the main sources of data. Government data has traditionally been accumulated in protected repositories and registries as public record and a matter of civil order. Consequently, over centuries, archived government data has had limited access via proprietary interfaces and often cumbersome fee-based procedures. With time going on, governments have started to look into the prospects of providing open access to their data repositories. The concept of Open Government Data (OGD) refers to government data defined as "data and information produced or commissioned by government or government controlled entities" that are opened up for use and re-use by public and private agents alike [3]. These data encompass both scientific data about the environment (from Earth Observation and other fields) and other public sector information, including diverse topics such as demographics, health and crime. In the current-used terminology, OGD does not include data that are subject to valid privacy, security or privilege limitations, as governed by other statutes. Government data sets are an interesting subset of open data because such subsets have already been collected for specific use, have been paid for by taxpayers, are relevant and offer value beyond what is captured from the originally intended use. When opened up, government data become a common, shared resource (i.e., public good) that is provided by the government. In our digital age, opening up government data to everyone puts power in the hands of citizens [3]. Making environmental data open is an effective method to guide the local agencies for policy-making and decision-making, and to increase the citizen participation for environmental protection.

Under the OGD circumstances, many private companies are set up to make business values in various sectors, such as business & legal services, data & technology, education, energy, finance & investment, lifestyle & consumer. It becomes common to observe that data itself has no value, the value comes from how data is used and analyzed in particular businesses. Specifically, private sectors, non-government organizations, businesses and APP developers in the environment/weather are encouraged to use the datasets already published to create economic, social values, and in particular environmental values (Sustainable Objectives). Many pioneering companies are using OGD as a key resource to generate values, for example, OGD-driven business models are proposed to evaluate how OGD is used as a key resource to understand the value generation process in e-business industry. The Open Data 500 Global Network is an international network of organizations that seek to study the use and impact of open data. Coordinated by the Governance Lab (GovLab) the OD500 Global Network enables participating organizations to analyze open data in their country in a manner that is both globally comparative and domestically specific. The OD500 Global Network starts from the assumption that only by mapping the use of open data within and across countries, can new approaches for understanding the economic and social impact of open government data be generated. The Open Data 500 U.S.A is the first comprehensive study of U.S. companies that use OGD to generate new business and develop new products and services [6]. Similarly, the Open Data 500 Korea, the Open Data 150 Canada, the Open Data 500 Australia, the Open Data 500 Mexico, the Open Data 500 Italy projects have been launched. Many companies from different sectors (e.g. healthcare, food and agriculture, finance etc.) are developed and generate value based on OGD from department of defense, department of energy etc. Our study show that around many environmental companies uses OGD to generate value, including the economic, social and environmental values.

Despite the potential significance of OGD, emphasized by an abundance of anecdotal evidence, we could not identify many studies on how OGD will contribute to value generations. To-date, the economic and social impact of open-data policies remains largely unclear, and there are scant empirical data available on the effects of the various policy approaches, thus leaving policy makers without the facts they need to assess and improve these policies. Some studies were conducted to illustrate how data are used as a key resource to enhance their core values [5,9,10]. However, more studies are needed to figure out how OGD is used as a key resource to generate values in specific domains, such as the environment/weather sector.

2 Research Motivation and Approach

Although a lot of anecdotal evidence has demonstrated the potential importance of OGD, it is still a mystery that how OGD will contribute to value generations in various sectors. Accordingly, our research on Open Government Data for Value Generation aims to address the question: How can use of OGD stimulate value generation in environment/weather companies? Building on a synthesis of the OGD literature and established theories of value generation, we develop a structural equation modelling (SEM)-based model to explore the causal relationship between OGD and value generation. The model is proposed to provide some guidance for the establishment of more companies to address more environmental issues. This research is planned in three phases. In the first phase the global open government data landscape in environment/weather sector is surveyed using secondary data from Open Data 500 projects. Accordingly, a conceptual SEM model is to be established for further research. In the second phase, questionnaires are designed and investigations conducted to collect data for model calculation and validation. In the third phase of this research, we plan to propose and implement appropriate OGD-driven companies in other countries to assist the governments for environmental protection.

This paper, however, only looks into the first phase of the planned research where it attempts to survey the landscape of OGD-driven environmental companies in USA, Australia and South Korea and establish a conceptual structural equation model for further validation.

3 OGD-Driven Companies in Environment/Weather Sector: The Current Situation

Many countries are entering the mainstream of open data movement. A number of open data benchmarks have been developed such as the World Bank's Open Data Readiness Assessment (ODRA), World Wide Web Foundation's Open Data Barometer (ODB, 2015), Open Knowledge Foundation Network's (2014) Open Data Index (ODI) and Capgemini Consulting's (2013) Open Data Economy (ODE), to name few global and widely used benchmarks. However, each of these benchmarks serves a different purpose and focus. Susha et al. (2014) suggest that ODB provides a more comprehensive perspective since it not only includes measures at various stages like readiness, implementation, and impact but also highlight the importance of involvement of major stakeholders and challenges throughout the open data process [9]. According to ODB (2016), a scaled score is an indication of how well a country is doing against other countries in getting the basics of open data readiness, implementation and impact [4]. According to the ODB (2014) technical handbook, When evaluating the ODB score, the openness of environmental data (D14 element) is considered based on the data on one or more of: carbon emissions, emission of pollutants (e.g. carbon monoxides, nitrogen oxides, particulate matter etc.), and deforestation. This study take 34 companies from US, Australia and South Korea into consideration due to the following two reasons. On one hand, the ODB rankings (as demonstrated in Fig. 2) show that US, Australia and South Korea are in the top ten countries in terms of open data readiness, implementation and impact. On the other hand, due to the high data openness in US, Australia and South Korea, many environmental companies participated in the Open Data 500 USA project, the Open Data 500 Australia project, and the Open Data 500 Korea project.

The 34 OGD-driven companies in environment/weather sector are surveyed and partially shown in Table 1, which present social impact, revenue (economic impact), and environmental impact of different companies. It is noted that the 34 companies use OGD from various official sources and their main purposes are to address severe environmental issues. The environmental issues to be addressed and other missions to be accomplished are also listed.

The aim of this study is to understand the causal relationship of why the OGD is beneficial for value generation from the environmental, social and economic aspects (Table 2).

Table 1. OGD-driven companies in environment/weather sector (Part 1)

No	Company name	City state country	Environmental impacts	Social impacts	Revenue sources	Data sources	Purposes	Websites
1	AccuWeather	State College_PA_US	Environmental protection, Climate change prevention	Environment and climate change	Subscriptions	National Weather Service, Monitoring stations	(1) Provide weather forecasts (2) Deliver enterprise solutions (3) Cooperate relationships between government weather agencies and the weather industry	http://www.accuweather.com
2	CoolClimate	Berkeley_CA_US	Environmentally sustainable development	Cleaner living environment	Data analysis for clients, Database licensing, User fees for web or mobile access	Multiple government open data sources (e.g. environmental records, land use information)	(1) Provide decision-making tools and programs (2) Design tailored climate solutions (3) Motivate low-carbon choices (4) Design climate actions and programs	http://coolclimate.berkeley.edu
3	Earth Networks	Germantown_MD_US	Extreme weather (e.g. tornadoes, cyclones), Climate change detection	Environment and climate change	Weather data analysis for organizations	Neighborhood-level sensors, National Oceanic and Atmospheric Administration, National Weather Service	(1) Operate weather observation, lightning detection and climate (green gas) networks (2) Enable enterprises and consumers to make informed decisions	www.earthnetworks.com
4	Environmental Data Resources	Milford_CT_US	Environmental problem report	Environment and climate change	Data collection outsources, Data management for clients	U.S. Environmental Protection Agency, Multiple government open data sources (e.g. environmental records, land use information)	(1) Provide robust data (2) Provide smarter workflow tools for data management	http://www.edrnet.com

Table 1. (*Continued*)

No	Company name	City state country	Environmental impacts	Social impacts	Revenue sources	Data sources	Purposes	Websites
5	Bass Coast Landcare Network Inc	Bass_ VIC_AU	Environmentally sustainable development	Citizen engagement and participation, educational opportunity, good governance	Contributions/ donations, Government contract, Membership fees, Philanthropic grants	Department of Environment Agencies, Local Governments, Organizations	(1) Support healthy and resilient ecosystems (2) Deliver sustainable agricultural and environmental management practices	http://www.basscoastlandcare.org.au
6	Enviro-dynamics Pty Ltd	Hobart_ TAS_AU	Environmental protection, Climate change prevention	Citizen engagement and participation, educational opportunity	Consulting	Department of Environment Australian Bureau of Statistics	(1) Provide environmental solutions (2) Help business steer through the green tape (3) Support government to deliver timely and sustainable environmental policy and practice outcomes (4) Provide agricultural and environmental data analysis	www.enviro-dynamics.com.au

Table 2. OGD-driven companies in environment/weather sector (Part 2)

No	Company name	City state country	Environmental impacts	Social impacts	Revenue sources	Data sources	Purposes	Websites
7	Parklands Albury Wodonga Ltd	Wodonga_VIC_AU	Environmental protection, Climate change prevention	Citizen engagement and participation, educational opportunity, good governance, public safety	Contributions/donations, Government contract, Philanthropic grants	New South Wales Government	(1) Tackle environmental challenges (2) Care for public land (3) Work together with businesses, neighbors and all levels of government	www.parklands-alburywodonga.org.au
8	Southern New England Landcare	Armidale_NSW_AU	Environmental protection, Climate change prevention	Citizen engagement and participation, educational opportunity, food access and supply	Contributions/donations, Membership fees, Philanthropic grants, Government grants	Agriculture, ABARE, Environment, Mapping, OEH threatened species	(1) Lead, connect and enable communities to meet sustainable living goals	http://snelcc.org.au/web3/
9	Ecobrain	Geumcheon-gu_Seoul_KR	Environmental protection, Climate change prevention	N/A	Consulting, Data analysis for clients	Korea Meteorological Administration	(1) Develop as an environment monitoring system (2) Predict expansion paths of pollutants from plants (3) Offer company-specific contributions	www.ecobrainkr.com

Table 2. (*Continued*)

No	Company name	City state country	Environmental impacts	Social impacts	Revenue sources	Data sources	Purposes	Websites
10	SBIS	Yeongdeungpo-gu_ Seoul_KR	Environmental protection, Climate change prevention	Citizen engagement and participation	Consulting, Data analysis for clients	Korea Meteorological Administration	(1) Provide reliable systems for customers	http://www.sbis.co.kr
11	Softworx	Nowon-gu_ Seoul_KR	Environmental protection, Climate change prevention	Consumer empowerment	Advertising, Software licensing	Korea Environment Corporation	(1) Offer information on air pollution and fine dust	https://play.google.com/store/apps/details?id=com.softworx.cai
12	w365	Gangnam-gu_ Seoul_KR	Environmental protection, Climate change prevention	Consumer empowerment	Consulting, Data analysis for clients	Korea Meteorological Administration	(1) Provide weather information	www.w365.com

Position	Country	Score	Readiness	Implementation	Impact
1	UK	100	100	100	100
2	USA	81.89	97	76	76
2	France	81.65	97	76	74
4	Canada	80.35	89	84	67
5	Denmark	76.62	77	77	78
6	New Zealand	76.35	87	62	87
7	Netherlands	75.13	90	69	70
8	Korea	71.19	95	64	58
9	Sweden	69.26	88	60	64
10	Australia	67.99	84	77	39
Average top 10		78.04	90.04	74.50	71.30

Fig. 1. Top ten countries in the ODB 3rd edition ranking (Source: 2016 ODB global report)

4 SEM Conceptual Model for Value Generation

OGD-driven companies in environment/weather sector apply OGD as a key resource in their business to generate sustainability values. Accordingly, it is necessary to reflect the causal relationship between OGD and the generated values. In the research, the SEM is proposed as a cause -and-effect diagram to illustrate how OGD is used to generate values.

4.1 Basic Structural Equation Modeling

Since Bentler's appeal to apply the technique to handle latent variables (i.e. unobserved variables) in psychological science, structural equation modeling (SEM) has become a quasi-routine and even indispensable statistical analysis approach in the social sciences. The emergence and development of SEM was regarded as an important statistical development in social sciences in recent decades and this "second generation" multivariate analysis method has been widely applied in theoretical explorations and empirical validations in many disciplines. Structural Equation Models with latent variables are extensively in measurement and hypothesis testing [1,2]. Compared with other statistical tools such as factor analysis and multivariate regression, SEM carries out factor analysis and path analysis simultaneously, since it can (1) measure and accommodate errors of manifest variables (i.e. observed variables); (2) represent ambiguous constructs in the form of latent variables (i.e. unobserved variables) by using several manifest variables; and (3) simultaneously estimate both causal relationships among latent variables and manifest variables. In addition, SEM can also provide group comparisons with a holistic model, resulting in much more vivid impressions than traditional ANOVA. SEM can also handle longitudinal designs when time lag variables are involved.

As introduced above, SEM describes and tests relationships between two kinds of variables – latent variables (LVs) and manifest variables (MVs). Latent variables cannot be observed directly due to their abstract character. In contrast, observed variables contain objective facts and easier to measure. Several observed variables can reflect one latent variable. As presented in Fig. 2, a structural equation model usually consists of two main components, a structural model and several measurement models. A simple measurement model includes a latent variable, a few associated observed variables and their corresponding measurement errors. The structural model consists of all LVs and their interrelationships. For model development purposes, some researches aim to validate their assumptions of a dimensional framework of one or several discriminant LVs, while others aim to elicit the causal relationship between the LVs. Confirmatory factor analysis (CFA) with correlating latent variables satisfies the former purpose, while these correlations need to be replaced by directional relationships for the latter.

Figure 4 provides a simple example of a structural equation model investigating the effect of LVs ξ_1, ξ_2 and ξ_3 on LVs η_1, η_2 and η_3, and where several MVs are used to represent the LVs. The MVs are shown in rectangles, the LVs in ellipses, measurement errors in circles and with arrows indicating the direction of the effects. If directional arrow between ξ and η is replaced by a correlation two-way arrow, the model is a CFA and its purpose is to test whether MVs can represent LVs well (i.e. convergent validity) and whether ξ and η are different (i.e. discriminant validity). The basic concepts and principles of SEM are now well established with the help of early explorations by researchers, structured textbooks, well developed soft programs (e.g. LISREL, EQS and AMOS).

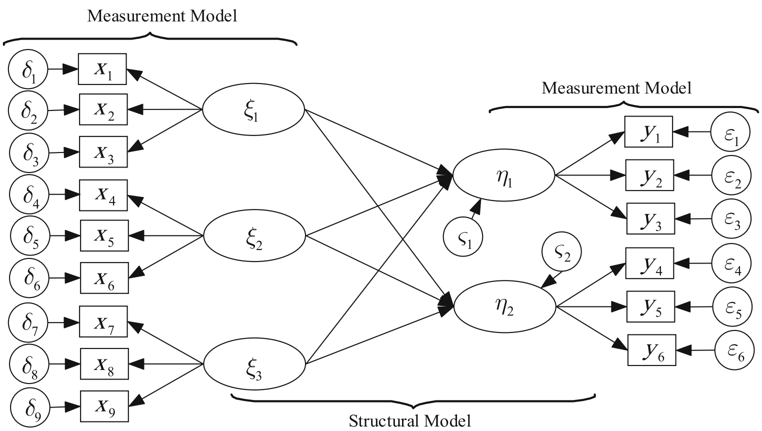


Fig. 2. A structural equation model diagram

4.2 Model Path Analysis

In order to establish a cause -and-effect diagram to illustrate how OGD is used to generate values, a path analysis is conducted, along with which we decide necessary latent variables and manifest variables based on a large amount of literature reviews.

(1) Constructs of OGD that are important in value creation
According to data-intensive science, data is prone to generate value through a value chain of big data, including value capture, value curation, value analysis and value visualization. To make sure that decision makers are able to produce information, knowledge or wisdom by utilizing various data analytics techniques and technologies. OGD is applied as a core resource for creating economic profits for themselves, bringing social values for the society and promoting environmental protection. In order to find out in what aspects OGD is important when firms use OGD to guide decision making or enhance transparency for value generation. The following features of data are essential for whether or not they are able to create values.

① Data quality & Integrity
OGD exposes data to public view in a way that allows for crowdsourced and quality control. Poor quality data can mislead for decision for environmental data-driven companies.

② Data disaggregation
OGD has many elements of being disaggregated and being highly accessible to many or even any party that wishes to use it. To some extent, disaggregated data can provide a better comparative picture of what works and help inform and promote evidence-based policy-making or decision making in environment/weather sector.

③ Data timeliness
OGD allows for private-sector and citizen contributions to open datasets that can make data more timely and relevant. In order to achieve real-time weather report, the data cycle has to match the decision cycle for better decision-making. Environmental data timeless in environment/weather sector is quite important for firms to make decisions, produce new products or provide new services.

④ Data transparency and openness
The promotion of OGD and tightened feedback loops between data users and data suppliers (namely government or government-governed entities). Publicly-funded datasets, as well as data on public spending, should be available to other public ministries or the general public. With high data transparency and openness, data should be made public in ways that encourage greater use and be complete, machine-readable, freely available for reuse without restrictions, and transparent about underlying assumptions. Environmental data transparency and openness in environment/weather sector is quite important for firms to make decisions, produce new products or provide new services.

⑤ Data usability and curation
Through OGD roundtables and similar events, data users can provide feedback on open datasets that makes them more accessible and useful. Data architecture should place great emphasis on user-centered design and user-friendly interfaces. Communities should be fostered to develop new tools that can translate raw data into something meaningful to a broader constituency of non-technical potential users.

⑥ Data protection and privacy
The challenge of opening government datasets while anonymizing sensitive information on health or environment sets a high bar for developing data privacy safeguards. Clear international norms and must be developed to protect data. The data protection is quite important when environmental companies utilize data in their daily business for weather forecast or policy suggestions.

⑦ Data governance and independence
The open data charter, open government partnership, and other international efforts are beginning to provide a governance structure for open data. Data quality should be protected and improved to ensure they are functionally autonomous, and independent of political influence.

⑧ Data resource and capacity
National statistical systems should be established that are capable of producing high quality statistics in line with global standards and expectations. Under these circumstances, people can make sure that the companies are able to provide reliable suggestions and conduct right measures for environmental protection.

(2) Constructs of value creation mechanisms
With rapid development of the society, data has become an important resource for creating values and we have entered the big data era. Most firms are utilizing big data for creating values and data-intensive science has become a new discipline. Data, in nature, has no value and data analysts aim to extract values, information, knowledge and wisdom by applying various techniques learning, data mining and technologies. It is the job of computer scientist to develop tools, but it is the practitioners' job to make decisions, provide new offerings or enhance transparency for creating values. The complementarity between the superior computational capabilities provided in the data analytics process and the higher judgmental capabilities of firm managers is quite important for realizing values from big data. Our study mainly consider three value creation mechanism based on literature reviews, and they are guiding decision makings, providing new offerings and enhancing transparency.

① Decision making
Decision making is quite important for firms to create values. By making better decisions based on data analytics, managers are able to achieve more economic or social values. Data-driven decision-making positively affects value through generation of new information, knowledge and wisdom.

② New offerings
Innovation is the source of value creation in Schumpeter's economic theory, bringing about novel combinations of resources, new production methods, as well as new products and services, which, in turn, lead to the transformation of markets and industries, thus increasing value. Data-driven innovation positively affects value through generation of new knowledge, new processes, services and products, and new businesses.

③ Transparency & accountability
Most definitions of transparency recognize the extent to which an entity reveals relevant information about its own decision processes, procedures, functioning and performance. However, opening access to chosen public documents does not necessarily contribute to a transparent government. By utilizing OGD for companies in environment/weather sector, it is important to promote the transparency and accountability of OGD from the environment sector.

(3) Constructs of values for OGD-driven companies in environment/weather sector
Two types of value are frequently discussed: economic value, defined as the worth of a good or service as determined by the market, and social value, which is created when resources, inputs, processes or policies are combined to generate improvements in the lives of individuals or society as a whole. The main difference is that our research emphasizes the environmental impacts of those companies from the environment/weather sector as the 34 companies make great contributions to environmental protection and disaster prevention. The environmental, economic and social values are combined to achieve the sustainability goal of the world, as shown in Fig. 3. The environmental impacts mainly include pollution reduction, natural resource conservation and climate change resilience.

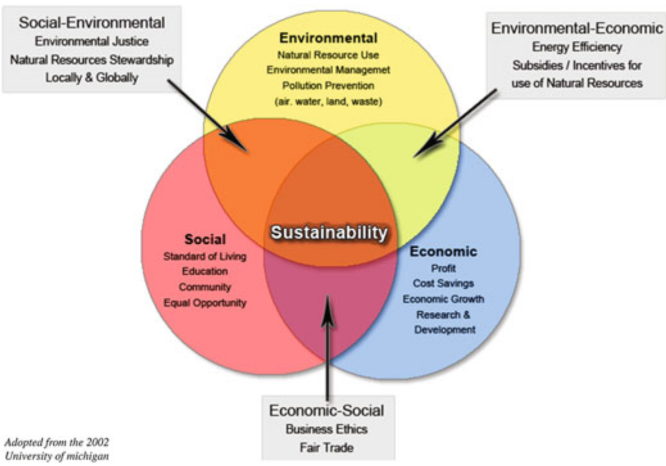


Fig. 3. Constructs of values

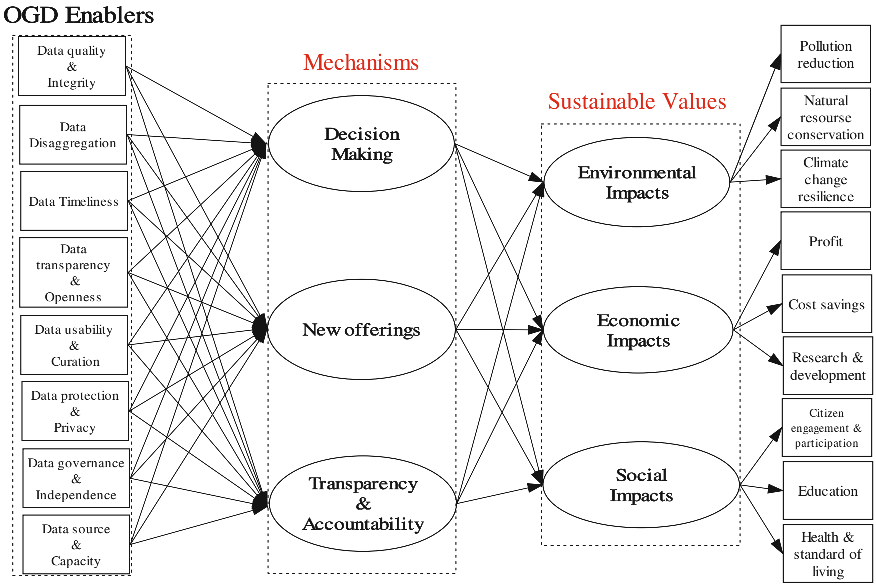


Fig. 4. Path analysis

The economic impacts are reflected in profits, cost savings, and research & development. The social impacts are categorized as citizen engagement & participation, education, and health & standard of living.

As previously discussed, the conceptual SEM model-based value generation methodology is established in Fig. 1.

5 Conclusions and Recommendations

Intensive research has been conducted to figure out the importance of big data or massive OGD, among which value chain analysis approach and value stream mapping method are extensively used. Both technical workers and management practitioners are key actors in finding out how big data or OGD contributes to value generations in various sectors. The beginning stage of our research on Open Government Data for Value Generation aims to address the issue of how OGD stimulates value generation in environment/weather companies. We attempt to analyze the causal relationship between OGD and sustainable values driven by OGD based on a structural equation model.

This paper, however, only looks into the first phase of the planned research where it attempts to survey the landscape of OGD-driven environmental companies in USA, Australia and South Korea and establish a conceptual structural equation model for further validation. Further research is needed to further enrich the current study.

Acknowledgements. This research was supported by the Program of China Scholarship Council (No. 201506240179).

References

1. Bagozzi RP, Yi Y (1988) On the evaluation of structural equation models. J Acad Mark Sci 16(1):74–94
2. Bagozzi RP, Yi Y (2012) Specification, evaluation, and interpretation of structural equation models. J Acad Mark Sci 40(1):8–34
3. Foundation OK (2016) Open data index (2016)
4. Fundation WWW (2016) Open Data Barometer, 3rd edn. (2016). http://www.opendatabarometer.org/report/about/method.html
5. Magalhaes G, Roseira C, Manley L (2014) Business models for open government data. In: Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance. ACM, pp 365–370
6. Network TOG (2016) Open data 500 U.S. (2016). http://www.opendata500.com/
7. (NRDC) NRDC (2016) Our stories (2015). https://www.nrdc.org/
8. Oceanic N, (NOAA) AA (2016) Climate data and report (2015). http://www.noaa.gov/
9. Susha I, Zuiderwijk A et al (2015) Benchmarks for evaluating the progress of open data adoption. Soc Sci Comput Rev 33(5):613–630
10. Zuiderwijk A, Janssen M (2013) Open data policies, their implementation and impact: a framework for comparison. Gov Inf Q 31(1):17–29

Impact of Management Information Systems Techniques on Quality Enhancement Cell's Report for Higher Education Commission of Pakistan

Faraz Ullah Khan and Asisf Kamran(✉)

Management and Social Science Department,
Institute of Business and Technology, Karachi, Pakistan
asifkamrankhan@gmail.com

Abstract. The way to achieve high quality education is through the quality affirmation. Higher education Commission has developed processes for the evaluation of the scholastics program's quality. The department of Quality Assurance has set up Quality Enhancement Cell which works under indicated Criterions. The point is to compose gathered information to produce useful information for QA clients. At present QEC is setting Self-Assessment Report through their predefined Performa and manual of models to help scholarly foundations. The fundamental issue of setting up this report is the utilization of manual strategies, which is the reason for delivering reports with repetitive errors and clearly, lumbering to oversee. The proposed solution for this issue is, as one may have seen in this time, fast move towards automation. MIS techniques will be applied by utilizing software development tools and database integration and broadly utilized reporting tools. Utilizing above mentioned tools and techniques will empower QA and clients to play out their job viably, reduce the repetitive information, save clients time and upgraded clarity. A more easy to understand Graphical User Interface helps client to accomplish more exact data for the decision making which will help them complete their tasks in impressively less time. This procedure have a major space for system enhancement system management more graphical reports more auxiliary changes by the assistance of software development tools.

Keywords: Self-Assessment Report · Higher education commission · Automation · Quality Enhancement Cell · Quality assurance · Management information systems

1 Introduction

Quality is a term which has always be confused with other concepts. The word quality has a Latin word "quails" its meaning is 'what sort of'. It has a range

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_30

of different meanings. Quality is something for you knowingly known. But there are things that are better than others. A high evaluation accorded to an educative process [11]. Where it has been demonstrated that, through the process, the students' educational development has been enhanced not only have they achieved the particular objectives set for the course but, in doing so, they have also fulfilled the general educational aims of autonomy of the ability to participate in reasoned discourse, of critical self-evaluation, and of coming to a proper awareness of the ultimate contingency of all thought and action [15].

Through above examinations, you will comprehend that esteem is undefined and setting focused. It stretches out from meaning "standard" to 'flawlessness'. Both are significantly settled in their different qualities operationalize in independent, institutions, and, the national practice will be visualize later. "Values can be justified in as a lowest 'threshold' by which the accuracy is judged" [4]. It is quite possible to have higher education institutions to have a common understanding in formulating to have teaching and research standards with high quality.

The administration measurement of value is most likely more likened to the instructive methodologies. We realize that not at all like physical products, administrations are vaporous to the degree that they can be expended just the length of the action or the methodology proceeds. Prominent issue of quality had been identified in the higher education in the Medium Term Development Framework (MTDF) [12]. Therefore, to enhance the quality of output and efficiency of the higher education learning systems, a mechanism of establishment of QECs has been developed by the Quality Assurance Committee to improve the standards of quality of higher education in an organized way with consistency across institutions. QEC prepared SAR to assess a systematic process of collecting, reviewing and using relevant quantitative/qualitative data and information from multiple and diverse sources about educational programs, for the purpose of improving student learning, and evaluating academic and learning standards. With a specific end goal to accomplish the subsequent reports the way toward finishing SAR is complicated [2]. The fundamental issue of setting up this report is the utilization of manual techniques, which is the reason for creating reports with repetitive errors and obviously, unwieldy to oversee. The answer for this issue is, as one may have seen in this time, quick move towards automation [13]. Applied MIS techniques by utilizing software development tools and database integration and generally utilized reporting tools. Utilizing above mentioned tools and techniques will empower QA and clients to perform their job effectively, lessen the repetitive data, save clients time and enhanced clarity [3].

1.1 Self-Assessment Report Manual by HEC (Document Prepared By: Dr. Abdul Raouf)

Criterion: Program Mission, Objectives and Outcomes

Each program must have a mission, measurable objectives and expected outcomes for graduates. Outcomes include competency and tasks graduates are expected to perform after completing the program. A strategic plan must be in place to achieve the program objectives. The extent to which these objectives are achieved through continuous assessment and improvements must be demonstrated [10].

Standard 1: The program must have documented measurable objectives that support Faculty/College and institution mission statements.

- Document institution, college and program mission statements.
- State program objectives. Program educational objectives are intended to be statements that describe the expected accomplishments of graduates during the first several years following graduation from the program.
- Describe how each objective is aligned with program, college and institution mission statements.
- Outline the main elements of the strategic plan to achieve the program mission and objectives.
- Provide for each objective how it was measured, when it was measured and improvements identified and made. Table 1 provides a format for program objectives assessment.

Table 1. Program objectives assessment

Objective	How measured measured	When measured	Improvement identified	Improvement made
1.				
2.	- do -			
3.	- do -			
4.	- do -			
5.	- do -			

Standard 2: The program must have documented outcomes for graduating students. It must be demonstrated that the outcomes support the program objectives and that graduating students are capable of performing these outcomes.

- Describe how the program outcomes support the program objectives. In Table 3 show the outcomes that are aligned with each objective..

- Describe the means for assessing the extent to which graduates are performing the stated program outcomes/learning objectives. This should be accomplished by the following (Table 2):

– Conducting a survey of graduating seniors every semester.
– Conduct a survey of alumni every two years.
– Conduct a survey of employers every two years.

Table 2. Outcomes versus objectives

Program objectives	Program outcomes			
	1	2	3	4
1.				
2.				
3.				

– Carefully designed questions asked during senior projects presentations. These questions should be related to program outcomes.
– Outcomes examinations.

The data obtained from the above sources should be analyzed and presented in the assessment report.

It is recommended that the above surveys should be conducted, summarized and added to the self-study assessment report. Departments should utilize the results of the surveys for improving the program as soon as they are available. An example follows:

1.2 Example (Program Objectives-Program Outcomes)

An example of program objectives and program outcomes is given below.

(1) Program Objectives (as developed by the department)

- Foundation
- Skills and Tools
- Awareness and Professional Ethics

① Objective 1
To provide students with a strong foundation in engineering sciences and design methodologies that emphasizes the application of the fundamental mathematical, scientific and engineering principles in the areas of engineering.

② Objective 2
To provide students with skills to enter the workplace well-prepared in the core competencies listed below:

- Design and modeling experience
- Open-ended problem solving ability
- Experimental and data analysis techniques
- Teamwork experience
- Oral written and multimedia communication skills
- Experience with contemporary computing systems and methodology

③ Objective 3
To provide students with knowledge relevant to engineering practice, including ethical, professional, social and global awareness, the impact of engineering on society, the importance of continuing education and lifelong learning in both technical and non-technical areas.

(2) Program Outcomes (as developed by the department)
Degree of skills and capabilities that will reflect on their performance as engineers:

- Students shall have an ability to apply knowledge of mathematics science and fundamental engineering to mechanical engineering problems.
- Students shall have an ability to identify, formulate and solve practical engineering problems.
- Students shall have an ability to design components, processes and systems to meet desired needs.
- Students shall have an ability to conduct engineering experiments to study different engineering systems, including various modes of operation, performance evaluation, properties of materials and manufacturing techniques, as well as to use laboratory instruments and computers to analyze and interpret data.
- Students shall have an ability to use modern tools, techniques, and skills necessary for practicing mechanical engineering including computational tools, statistical techniques, and instrumentation.
- Students shall have an ability to work in a professional engineering environment, and to understand the associated economical considerations.
- Students shall have an ability to communicate effectively in written, oral, and graphical forms, including the use of professional quality visual aids.
- Students shall have an ability to work effectively in teams including multidisciplinary teams to solve engineering problems relevant to their field.
- Students shall have an understanding of the professional and ethical responsibilities of engineers.
- Students shall have an understanding of the impact of engineering on society and environment.
- Students shall have recognition of the need and an ability to engage in life long learning of engineering.
- The program outcomes are the by products of the program objectives and are interrelated. An example of interrelation between the program objectives and the program outcomes is shown in the following table.

Standard 3: The results of program's assessment and the extent to which they are used to improve the program must be documented.

- Describe the actions taken based on the results of periodic assessments.
- Describe major future program improvements plans based on recent assessments.

Table 3. Relationship between Program objectives and Program outcomes

Program objectives	Program outcomes										
	◎	2	3	4	5	6	7	8	9	◎0	◎◎
1	◎	◎	◎	.	◎	.	.	.	.	.	.
2a	◎	◎	◎	◎	.	◎	.	.	.	.	.
2b	◎	◎	◎	◎	.	◎	.	.	.	.	.
2c	◎	.	.	◎	◎	◎	.	.	.	.	.
2d	.	.	.	.	.	◎	.	◎	.	.	.
2e	.	.	.	.	.	◎	◎	◎	.	.	.
2f	.	.	.	.	◎	◎	.	.	.	.	.
3	.	.	.	.	.	◎	.	◎	◎	◎	◎

Legend: ◎ Denotes substantial contribution to the objective and ⊖ denotes moderate contribution to the objective. . Denotes no contribution to the objective.

- List strengths and weaknesses of the program.
- List significant future development plans for the program.

Standard 4: The department must assess its overall performance periodically using quantifiable measures.

- Present students enrolment (undergraduate and graduate) during the last three years indicating percentages of honor students, student faculty ratio, average graduating grade point average per semester, average time for completing the undergraduate program and attrition rate.
- Indicate percentage of employers that are strongly satisfied with the performance of the department's graduates. Use employer's survey.
- Indicate the median/average student evaluation for all courses and the % of faculty awarded excellence in teaching award.
- Present performance measures for research activities. These include journal publications, funded projects, and conference publications per faculty per year and indicate the % of faculty awarded excellence in research award.
- Present performance measures for community services. This may include number of short courses per year, workshops and seminars organized.
- Indicate faculty and students satisfaction regarding the administrative services offered by the department. Use faculty and students surveys.

2 Problem Statement

The main problem of preparing this report is the use of manual methods, which is the cause of producing reports with repetitive errors and obviously, cumbersome to manage. The proposed solution to this problem is, as one may have observed in this era, rapid move towards automation.

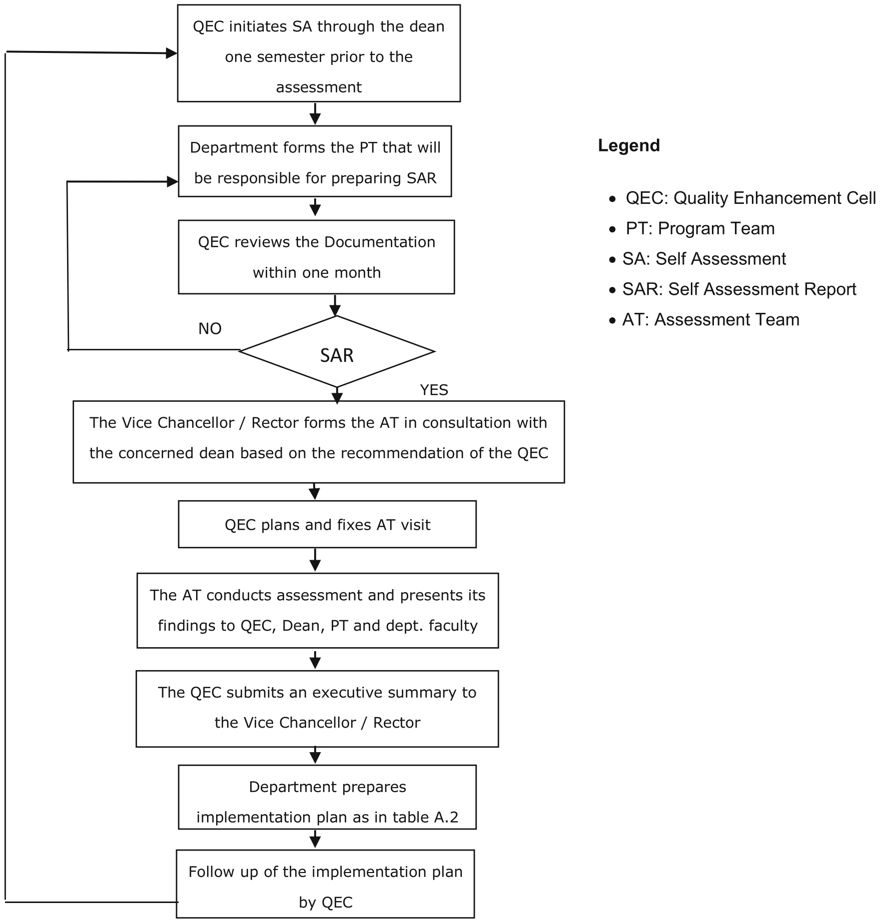


Fig. 1. Self-Assessment procedure

3 Method of Research

This research is aimed that of the process of Self-Assessment Report gets automated by using technology, for example through user friendly software then this will help in enhance the process of QA Reporting for HEC. Therefore this research is a Descriptive Research in which the method of research is Observational Research. Since the respondents in this research work in a natural environment, this method is convenient in assessing multiple users performing their jobs.

Using detailed investigation methodology adds a lot for the investigation, for instance:

- Detailed research reveals difficulties as well as excessive ailments to ensure that remedial procedures may be instituted. This might assistance in presenting ideas for predicaments exactly where progress may be required.
- Detailed exploration facilitates the particular prediction of the future by results about existing disorders, effects and by persons in the direction of a particular matter.
- Detailed exploration provides a much better and much deeper comprehension of any trend by an in-depth study from the trend.

3.1 Research Methodology

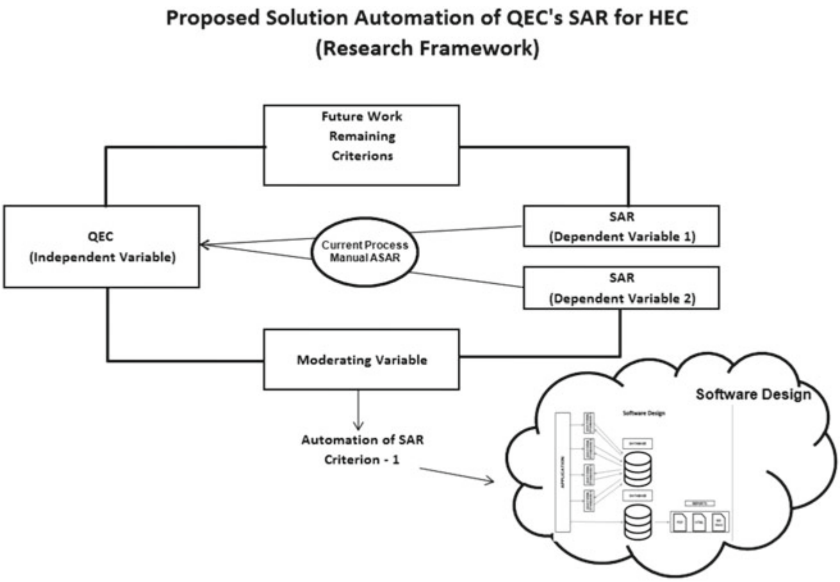


Fig. 2. Proposed solution automation of QEC's SAR for HEC

3.2 Design of Research

In view of these variables and hypothesis, the Research Design followed for this thesis is Qualitative Research Design. Reasons for choosing this design:

- This design helps in identifying & controlling extraneous factors also.
- This design is efficient in using available resources.

3.3 Techniques & Instrument of Research

Research instruments included a single List of questions. The set of questions was structured. Various factors relevant to the study have been integrated from the set of questions. There were 20 issues in which concerned numerous qualities in addition to considerations. The set of questions was made in view from the study ambitions. The options from the set of questions derive from Contract by making use of Liker Range.

- Strongly Agree
- Agree
- Disagree

3.4 Target Population

The sample of the SAR user is selected for the research, those who are working for QEC. Those individual SAR users who are preparing their report by manual method. On a whole we have taken almost 61 responses from different institutions.

3.5 Data Analysis

The purpose of the research is to analyze the aspects which are affecting the adoption of Automation of QEC's SAR. The researcher selected the questionnaire for collecting the responses. The average graph of respondents is showing high requirement for the automation of SAR.

3.6 Research Methodology

The representation of the operation sequence uses job-based encoding [13], and the length of the chromosome equals the total number of operations. The job number denotes the operation of each job, and the $l-th$ occurrence of a job number refers to the $l-th$ operation in the sequence of this job. Figure 1 presents an example considering a rescheduling with a 4 jobs and 3 machines.

3.7 Impact of Research Question

The study will apply the following impacts, when the SAR will automate.

- Time saving
- Process
- Error Reduction

Figure 3(a) shows the significance of Automation of Self Assessment Report by high requirement of respondents. Figure 3(b) shows the significance of process Improvement viewed by respondents. Figure 3(c) shows the significance of Time saving can be viewed by the positive responses. Figure 3(d) shows the magnitude of error reduction is shown by the positive responses.

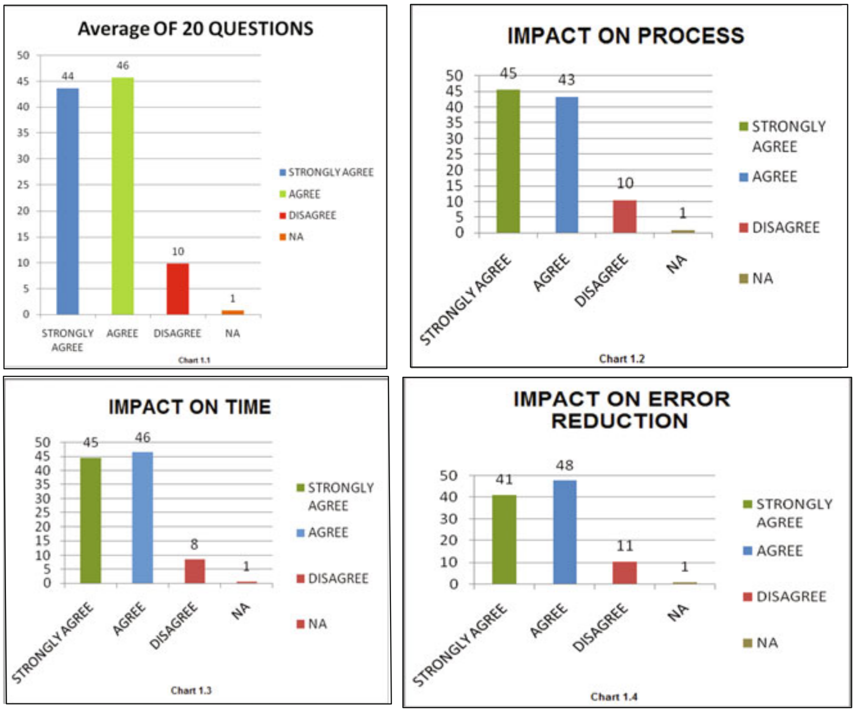


Fig. 3. Questionnaire graphical analysis

4 Conceptual Model

The above given diagram is a conceptual model of all entities participating in the process of making SAR. This model only shows how the whole process works currently without having being automated. From this diagram one can start to figure out how the software can be evolved in the system to provide or achieve the outcomes which are supposed to be achieved.

The process starts when an institution wishes to approve its program by HEC. QEC personal sends a request to HEC which in return supplies 'SAR Manual' in order to prepare SAR Report. QEC user inserts MISSION OBJECTIVES & OUTCOMES based on the CRITERION-1 in Self Assessment Report Manual. Finally report is generated as per the format of HEC.

4.1 Waterfall Models

There are different models used for software life cycle the one which is suitable in SAR is Waterfall Model. As this model distinguishes the whole process into distinct phases.

"A software requirements specification (SRS) is a description of a software system to be developed, laying out functional and non-functional

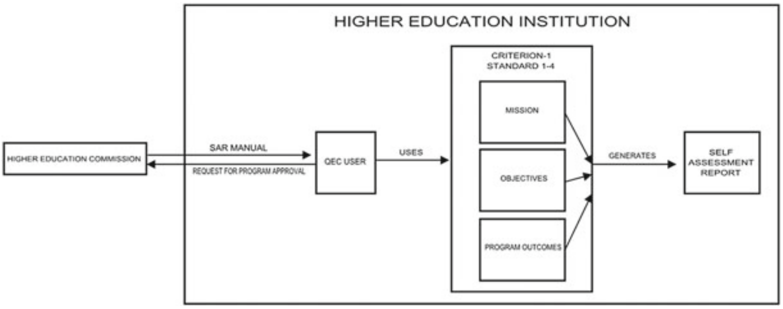


Fig. 4. Conceptual model

requirements" [8]. This model is used as a frame work or it provides a specified format to develop any software in a predefined phases. It starts from gathering requirements from users and then moves on to produce software design. Later this design will be implemented after unit testing. The last two phases may not need very much consideration in SAR.

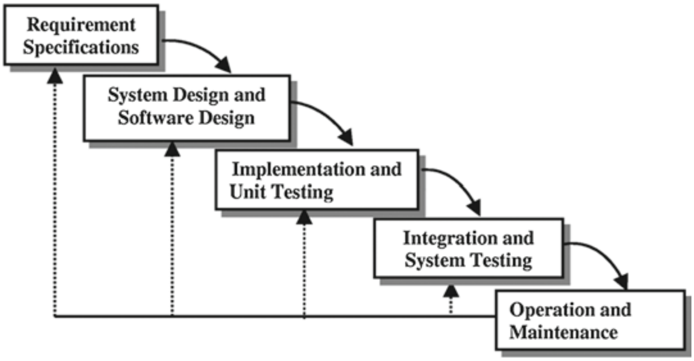


Fig. 5. Waterfall models

4.2 Languages

For the above purpose, C# & VB Languages have been selected to process SAR Report. C# & VB is usually a multi-paradigm programming language encompassing strong typing, essential, declarative, sensible, universal, object-oriented (class-based), and also component-oriented programming disciplines. C# & VB is one of the programming languages designed for the Common Language Infrastructure.

4.3 Programming Methodology

Object-Oriented Programming has the following advantages over conventional approaches [7]:

Provides a clear modular structure for programs which makes it good for defining abstract data types where implementation details are hidden and the unit has a clearly defined interface.

- Makes it easy to maintain and modify existing code as new objects can be created with small differences to existing ones.
- Provides a good framework for code libraries where supplied software components can be easily adapted and modified by the programmer. This is particularly useful for developing graphical user interfaces [1].

4.4 Platform

Visual Studio is a complete set of development tools for building C# & VB applications, XML Web Services, desktop applications, and mobile applications. Visual Basic, Visual C#, and Visual C++ all use the same integrated development environment (IDE), which enables tool sharing and eases the creation of mixed-language solutions [14].

4.5 Database

A Relational database is a computerized database whose association is focused around the Relational model of information, as proposed by E.f. Codd in 1970. This model arranges information into one or more tables (or "relations") of columns and sections, with a Primary key for each one line [6]. By and large, every element sort portrayed in a database has its own particular table, the lines speaking to occurrences of that substance and the segments speaking to the characteristic qualities depicting each one occurrence (Education Indicators [9]. Since each one column in a table has its own particular interesting key, pushes in different tables that are identified with it can be connected to it by putting away the first line's one of a kind key as a characteristic of the optional line (where it is known as a "remote key"). Codd [6] demonstrated that information connections of subjective unpredictability can be spoken to utilizing this straightforward set of ideas. The different programming frameworks used to keep up social databases are known as Relational Database Management Systems (RDBMS).

Essentially all Relational database frameworks use SQL (Structured Query Language) as the dialect for questioning and keeping up the database.

A good choice would be "Microsoft SQL Server is a relational database management system developed by Microsoft. As a database, it is a software product whose primary function is to store and retrieve data as requested by other software applications, be it those on the same computer or those running on another computer across a network (including the Internet)" [17].

Fig. 6. Types of reporting tool

4.6 Reporting Tool

The Visual Studio Report Designer provides a user-friendly interface for creating robust reports that include data from multiple types of data sources [5] (Fig. 6).

'Visual Studio' reports let you slice and dice your data and present it in detail or summary form regardless of how the data is stored or sorted in the underlying tables. It offers a great deal of power and flexibility to analyze and present results [16].

4.7 Software Architecture

User authentication and password protection Managers and QEC directors given the authorization to change all the aspects of SAR report the director is working with are also given the authorization to access the whole report but with the exception of few section which are in read only mode.

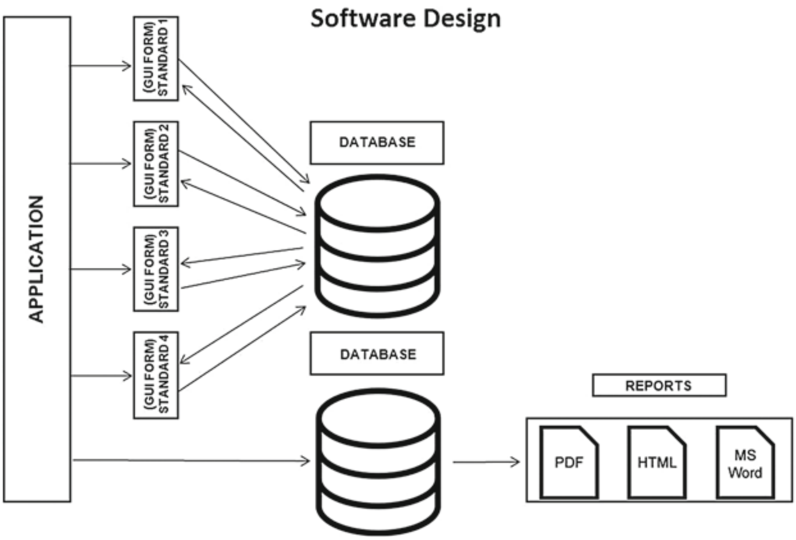


Fig. 7. Software design

4.8 Software Design

Is the process by which an agent creates a specification of a software artifact, intended to accomplish goals, using a set of primitive components and subject to constraints. Software design usually involves problem solving and planning a software solution [17]. The figure explains main components of ESAR Design. Application interacts with the four graphical user interfaces (GUI) sequentially. GUI component establishes two way communication with database components (that comprises related tables) once all the required data is stored in database, user initiates the process of generating report (an ultimate components of ESAR) this final result will also be saved (Fig. 7).

5 Result

- Figure 2 illustrates the research framework and demonstrate the function of moderate variable.
- Figure 4 shows the process of Self-Assessment Report through conceptual model.
- Figure 5 demonstrates the software design of Self-Assessment Report.
- TIME: the process of ESAR will significantly save the user's time.
- MIS Techniques: the above advantages of results may well be achieved MIS techniques.
- REDUCTION OF REPEATION ERRORS: using RDBMS.
- FLEXIBILITY IN FUTURE MODIFICATION: can be achieved using OOP techniques.

6 Conclusions

Quality assurance is a portal to "High quality education". With a specific end goal to evaluate the programs quality Higher Education Commission" has created predefine Criterions forms. The department of Quality Enhancement Cell which works under those predefined Criterion. This Criterion is met utilizing frames with applicable information. Thusly, this will bring about an enhanced interaction with QA clients utilizing E-SAR. One of the other major advantages of it will be as better organized report. In this manner, the real issue for QA clients of following QEC procedures can extensively be all around encouraged utilizing the proposed solution. Infect, repetitive errors are better managed with it. A well designed application is most suitable to accomplish the proposed results for future work. In future, alternative design technique of HCI can be utilized to furnish client with more upgraded collaboration in SAR making.

In order to discuss the process of QEC's SAR, we clearly observe that the whole process is currently based on manual work. There is a dire need to automate SAR process which reflects from the survey data analysis, collected from the QEC users. The remaining Criterion of SAR can easily be considered by applying MIS technology and software development tools in order to automate.

References

1. Afaf (2016) Advantages of OOP. https://www.researchgate.net/post/What_is_adventages_of_Object_Oriented_Programming
2. Alvi NA, Alam A (2004) Pakistan institute of quality control
3. Arshad M (2003) Attitude of teachers of higher education towards their profession. Ph.D. thesis, Allama Iqbal Open University, Islamabad
4. Ashcroft K (2005) Criteria standards and quality. https://books.google.com.pk/books?isbn=113571990X
5. Code EF (1990) Keys and referential integrity. https://www.seas.upenn.edu:codeblab.com/wp-content/uploads/2009/12/rmdb-codd.pdf
6. Code EF (1990) The relational model for database management: verison 2[M]. Addison-Wesley Longman Publishing co. Inc
7. Coulter N (2015) Advantages and disadvantages of OOP
8. ISO/IEEE 830-1998 (1998) Recommended practice for software requirements specifications. doi:10.1109/IEEESTD.1998.88286
9. Indicators E (2015) Expenditure on education as % of total government expenditure (%). http://data.worldbank.org/indicator/SE.XPD.TOTL.GB.ZS
10. Institute BS (2015) Learning and growth perspective. http://balancedscorecard.org/Learning-and-Growth-Perspective
11. James Hutton DP (2010) Corporate governance-good governance in the university context. http://www.minterellison.com/Pub/NL/201003_HEFd/
12. Khan MA, Usman M (2015) Education quality and learning outcomes in higher education institutions in pakistan. Springer, Singapore, pp 449–463
13. Michal Pietrzak JP (2015) The application of the balanced scorecard (bsc) in the higher education setting of a polish university. Online J Appl Knowl Manag 3(1):151–164
14. Microsoft (2008) Introducing visual studio. https://msdn.microsoft.com/en-us/library/fx6bk1f4(v=vs.90).aspx
15. Mishra S (2007) Concept of quality. http://oasis.col.org/bitstream/handle/11599/101/QAHE_Intro.pdf?sequence=1
16. Ralph P, Wand Y (2017) A proposal for a formal definition of the design concept. In: Lyytinen K, Loucopoulos P et al (eds) Design Requirements Workshop. LNBIP, vol 14. Springer, Heidelberg, pp 103–136
17. WikiPedia (2008) Microsoft sql server. https://en.wikipedia.org/wiki/Microsoft_SQL_Server

A Priority-Based Genetic Representations for Bicriteria Network Design Optimizations

Lin Lin[1,2,3(✉)], Jingsong Zhao[1], Sun Lu[1], and Mitsuo Gen[2]

[1] Dalian University of Technology, Dalian, People's Republic of China
lin@dlut.edu.cn
[2] Fuzzy Logic Systems Institute, Tokyo, Japan
[3] Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian, People's Republic of China

Abstract. Network design is one of the most important and most frequently encountered classes of optimization problems. It is a combinatory field in combinatorial optimization and graph theory. A lot of optimization problems in network design arose directly from everyday practice in engineering and management. Furthermore, network design problems are also important for complexity theory, an area in the common intersection of mathematics and theoretical computer science which deals with the analysis of algorithms. Recent advances in evolutionary algorithms (EAs) are interested to solve such practical network problems. However, various network optimization problems typically cannot be solved analytically. Usually we must design the different algorithm for the different type of network optimization problem depending on the characteristics of the problem. In this paper, we investigate the recent related researches, design and validate effective priority-based genetic representations for the typical network models, such as shortest path models (node selection and sequencing), spanning tree models (arc selection) and maximum flow models (arc selection and flow assignment) etc., that these models covering the most features of network optimization problems. Thereby validate that EA approaches can be effectively and widely used in network design optimization.

Keywords: Evolutionary Algorithm (EA) · Shortest path model · Spanning tree model · Maximum flow model · Bicriteria network design

1 Introduction

Many real-world problems from operations research (OR)/management science (MS) are very complex in nature and quite hard to solve by conventional optimization techniques. Since 1960s, there has been being an increasing interest in imitating living beings to solve such kinds of hard optimization problems. Simulating natural evolutionary process of human beings results in stochastic optimization techniques called evolutionary algorithms (EAs) that can often outperform conventional optimization methods when applied to difficult real-world

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_31

problems. EAs mostly involve metaheuristic optimization algorithms such as: genetic algorithm (GA) [18,21], evolutionary programming (EP) [12], evolution strategies (ES) [36,39], genetic programming (GP) [27,28], learning classifier systems (LCS) [22], swarm intelligence (comprising ant colony optimization (ACO) [9] and particle swarm optimization (PSO) [25,26]. Among them, genetic algorithms are perhaps the most widely known type of evolutionary algorithms today and load more stuff again.

Network design is one of the most important and most frequently encountered classes of optimization problems [2]. It is a combinatory field in graph theory and combinatorial optimization. A lot of optimization problems in network design arose directly from everyday practice in engineering and management: determining shortest or most reliable paths in traffic or communication networks, maximal or compatible flows, or shortest tours; planning connections in traffic networks; coordinating projects; and solving supply and demand problems. The most of network design problems where even one cost measure must be minimized are often NP-hard [14]. However, in practical applications, it is often the case that the network to be built is required to multiobjective. In this paper, we introduce three kinds of major bicriteria network design models. These models are the core multi-criteria network design models and extensively used in practical applications. (1) Bicriteria shortest path (bSP) model is one of the basic multi-criteria network design problems. It is desired to find a diameter-constrained path between two specified nodes with minimizing two cost functions. Hansen presented the first bSP model at 1979 [19]. Recently, Hao and Gen [20] examined an effective multiobjective EDA for bi-criteria stochastic job-shop scheduling problem; Shokrollahpour and Zandieh [47] presented A novel imperialist competitive algorithm for bi-criteria scheduling of the assembly flowshop problem. (2) Bicriteria spanning tree (bST) model play a central role within the field of multi-criteria network modes. It is desired to find a subset of arcs which is a tree and connects all the nodes together with minimizing two cost functions. Craveirinha et al. [4] presented a bi-criteria minimum spanning tree routing model for MPLS/overlay networks, and Steiner [41] Computing all efficient solutions of the biobjective minimum spanning tree problem.

2 Multiobjective Genetic Algorithms

Optimization deals with the problems of seeking solutions over a set of possible choices to optimize certain criteria. If there is only one criterion to be taken into consideration, it becomes to single objective optimization problems, which have been extensively studied for the past 50 years. If there are more than one criterion which must be treated simultaneously, we have multiple objective optimization problems [1,6,42]. Multiple objective problems arise in the design, modeling, and planning of many complex real systems in the areas of industrial production, urban transportation, capital budgeting, forest management, reservoir management, layout and landscaping of new cities, energy distribution, etc. It is easy to find that almost every important real world decision problem involves multiple and conflicting objectives which need to be tackled while respecting various constraints, leading to overwhelming problem complexity. The multiple objective optimization problems have been receiving growing interest from researchers with various background since early 1960 [23]. There are a number of scholars who have made significant contributions to the problem. Among them, Pareto is perhaps one of the most recognized pioneers in the field [11]. Recently, GAs have been received considerable attention as a novel approach to multiobjective optimization problems, resulting in a fresh body of research and applications known as evolutionary multiobjective optimization (EMO).

2.1 Fitness Assignment Mechanism

GAs are essentially a kind of meta-strategy methods. When applying the GAs to solve a given problem, it is necessary to refine upon each of the major components of GAs, such as encoding methods, recombination operators, fitness assignment, selection operators, constraints handling, and so on, in order to obtain a best solution to the given problem. Because the multiobjective optimization problems are the natural extensions of constrained and combinatorial optimization problems, so many useful methods based on GAs developed during the past two decades. One of special issues in the multiobjective optimization problems is fitness assignment mechanism. Since the 1980s, several fitness assignment mechanisms have been proposed and applied in multiobjective optimization problems [7,8,13,15,16,24,38,40,49]. Although most fitness assignment mechanisms are just different approach and suitable to different cases of multiobjective optimization problems, in order to understanding the development of multiobjective GAs, we classify algorithms according to proposed years of different approaches:

2.2 Performance Measures

Let S_j be a solution set $(j = 1, 2, \cdots, J)$. In order to evaluate the efficiency of the different fitness assignment approaches, we have to explicitly define measures evaluating closeness of S_j from a known set of the Pareto-optimal set S*. For example, the following common 3 measures are considered that already used in different moGA studies [24]. They provide a good estimate of convergence if a reference set for S* (i.e., the Pareto optimal solution set or a near Pareto optimal solution set) is chosen.

(1) Number of Obtained Solutions $|S_j|$

Evaluate each solution set depend on the number of obtained solutions.

(2) Ratio of Nondominated Solutions $R_{\mathrm{NDS}}\ (S_j)$

This measure simply counts the number of solutions which are members of the Pareto-optimal set S^*. The $R_{\mathrm{NDS}}(S_j)$ measure can be written as follows:

$$R_{\mathrm{NDS}}(S_j) = \frac{|S_j - \{x \in S_j\, |\exists r \in S^* : r \prec x\}|}{|S_j|},$$

where $r \prec x$ means that the solution x is dominated by the solution r. The $R_{\mathrm{NDS}}(S_j)$ =1 means all solutions are members of the Pareto-optimal set S^*, and $R_{\mathrm{NDS}}(S_j) = 0$ means no solution is a member of the S^*. It is an important measure that although the number of obtained solutions $|S_j|$ is large, if that the ratio of nondominated solutions $R_{\mathrm{NDS}}(S_j)$ is 0, it may be the worst result. The difficulty with the above measures is that although a member of S_j is Pareto-optimal, if that solution does not exist in S^*, it may be not counted in $R_{\mathrm{NDS}}(S_j)$ as a non-Pareto-optimal solution. Thus, it is essential that a large set for S^* is necessary in the above equations.

(3) Average Distance D1R(S_j)

Instead of finding whether a solution of S_j belongs to the set S^* or not, this measure finds an average distance of the solutions of S_j from S^*, as follows:

$$D1_R(S_j) = \frac{1}{|S^*|} \sum_{r \in S*} \min\{d_{rx} \,|x \in S_j\},$$

where d_{rx} is the distance between a current solution x and a reference solution r in the 2-dimensional normalized objective space. f_i means the objective function for each objective $i = 1, 2, \cdots, q$.

$$d_{rx} = \sqrt{\sum_{i=1}^{q} (f_i(r) - f_i(x))^2}.$$

The smaller the value of $D1_R(S_j)$ is, the better the solution set S_j is. This measure explicitly computes a measure of the closeness of a solution set S_j from the set S^*.

(4) Reference Set S^*

For making a large number of solutions in the reference set S*, the first step calculates the solution sets with special GA parameter settings and much long computation time by each approach which used in comparison experiments, the second step combine these solution sets to calculate the reference set S^*. In the future, a combination of small but reasonable GA parameter settings for comparison experiments will be conducted. Thus, ensure the effectiveness of the reference set S^*.

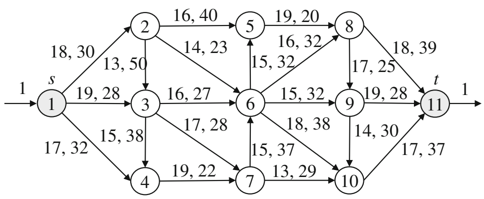


Fig. 1. A simple network with 11-node/ 22-edge

3 Bicriteria Shortest Path (bSP) Model

Routing problem is one of the important research issues in communication network fields. As B-ISDN standards and ATM networks emerge, modern networks are expected to provide a wide range of services and guarantee various end-to-end Quality of Service (QoS). Traditionally, the routing problem has been a single objective problem having as goal the minimization of either the total distance or travel time. However, in practical applications of communication network, create an effective route consider not only cost or travel time, but also requirements of social and economic environment. In other words, it is necessary to take into account that many real-world problems are multiobjective in nature. The objective functions related to cost, time, accessibility, environmental impact, reliability and risk are appropriated for selecting the most satisfactory ("best compromise") route in many communication network optimization problems [3].

3.1 Mathematical Formulation

Let $G = (N, A)$ be a directed network, consisting of a finite set of nodes $N = 1, 2, \cdots, n$ and a set of directed arcs $A = \{(i, j), (k, l), \cdots, (s, t)\}$ joining m pairs of nodes in N. Arc (i, j) is said to be incident with nodes i and j, and is directed from node i to node j. Suppose that each arc (i, j) has assigned to it nonnegative numbers c_{ij}, the cost of (i, j) and d_{ij}, the transmission delay of (i, j). Let x_{ij} is a 0-1 Decision variables the link on an arc $(i, j) \epsilon$ A. The bSP problem is formulated as follows, in which the objectives are to minimizing total cost z_1 and minimizing transmission delay z_2 form source node 1 to sink node n.

$$\min z_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{1}$$

$$\min z_2 = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_{ij} \tag{2}$$

$$s.t. \sum_{j=1}^{n} x_{ij} - \sum_{k=1}^{n} x_{ki} \begin{cases} = 1, & i = 1 \\ = 0, & i = 2, 3, \cdots, n-1 \\ = -1, & i = n \end{cases} \tag{3}$$

$$x_{ij} = 0 \text{ or } 1, \quad (i, j = 1, 2, \cdots, n), \tag{4}$$

where constraint (3), a conservation law is observed at each of the nodes other than s or t. That is, what goes out of node i, $\sum_{i=1} x_{ij}$ must be equal to what comes in $\sum_{k=1} x_{ki}$.

3.2 GA Approaches for bSP Model

How to encode a solution of the network design problem into a chromosome is a key issue for GAs. In Holland's work, encoding is carried out using binary strings. For many GA applications, especially for the problems from network design problems, the simple approach of GA was difficult to apply directly. There two special difficulties by using GAs for creating a path: (1) different path contains variable number of nodes; (2) a random sequence of nodes usually does not correspond to a path.

Zhang et al. [48] extended this variable-length chromosome for solving the SPR problem. But crossover may generate infeasible chromosomes that generating loops in the routing paths. It must be checked that none of the chromosomes is infeasible at each generation, and is not suitable for large networks or unacceptable high computational complexity for real-time communications involving rapidly changing network topologies. An example of generated variable-length chromosome and its decoded path are shown in Fig. 2(a) and (c), respectively for the directed network shown in Fig. 1.

We proposed a priority-based encoding method. As it is known, a gene in a chromosome is characterized by two factors: locus, i.e., the position of gene located within the structure of chromosome, and allele, i.e., the value the gene takes. In this encoding method, the position of a gene is used to represent node ID and its value is used to represent the priority of the node for constructing a path among candidates. A path can be uniquely determined from this encoding. An example of generated priority-based chromosome is shown in Fig. 2(b). At the beginning, we try to find a node for the position next to source node 1. Nodes 2, 3 and 4 are eligible for the position, which can be easily fixed according to adjacent relation among nodes. The priorities of them are 1, 10 and 3, respectively. The node 3 has the highest priority and is put into the path. The possible nodes next to node 3 are nodes 4, 6 and 7. Because node 6 has the largest priority value, it is put into the path. Then we form the set of nodes available for next position and select the one with the highest priority among them. Repeat these steps until we obtain a complete path, (1-3-6-5-8-11). Considering the characteristic of priority-based chromosome, they proposed a new crossover operator, called weight mapping crossover (WMX) and adopted insertion mutation and immigration operators.

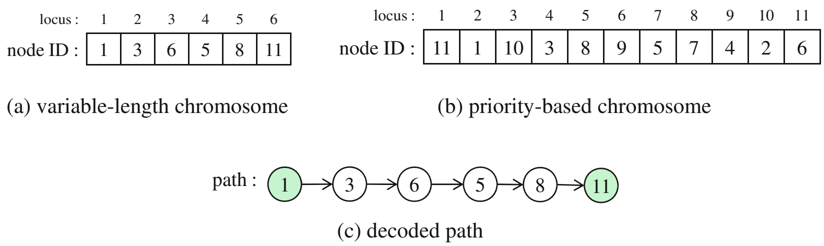


Fig. 2. An example of chromosome and its decoded path

3.3 Experiments and Discussion

In this subsection, the first experiment examines the different chromosome representations. We compare priority-based encoding with Ahn and Ramakrishna's variable-length encoding. Then to validate the multiobjective evolutionary approaches for solving bSPR problems, we compare our i-awGA with 3 different fitness assignment approaches (spEA, nsGA II and rwGA). All test network topologies were constructed by Beasley and Christofides [31]. Costs and delay are represented as random variables depend on the distribution functions.
(1) Different Encoding Methods
In order to evaluate the results of each test, we use the single objective by minimizing total cost, and combine average of the best solutions (ABS). In addition, we demonstrate the difference among the quality of solutions obtained by various GA parameter settings and our auto-tuning strategy. There are 3 kinds of different GA parameter settings:

Para 1: popSize = 10, $p_c = 0.3$, $p_M = 0.7$, $p_I = 0.30$
Para 2: popSize = 10, $p_c = 0.5$, $p_M = 0.5$, $p_I = 0.30$
Para 3: popSize = 10, $p_c = 0.7$, $p_M = 0.3$, $p_I = 0.30$.

In addition, two different stopping criteria are employed. One of them is the number of maximum generations, maxGen = 1000. Another stopping criteria is $T = 200$. That is, if the best solution is not improved until successive 200 generations, the algorithm will be stopped.

Table 1 shows the ABS of 50 runs by different GA parameter settings with different genetic representations respectively. As depicted in Table 1, most results of ABS of 50 runs by priGA with auto-tuning operator proposed are better than each of the other combinations, except to the test 4, test 6 and test 8.
(2) Different Fitness Assignment
In the second experimental study, we demonstrate the performance comparisons of multiobjective GAs for solving bSP problems by different fitness assignment approaches, there are spEA, nsGA II, rwGA and i-awGA. In each GA approach, priority-based encoding was used, and arithmetical crossover, swap mutation, immigration and auto-tuning operators were used as genetic operators.

As depicted in Table 2, most results of ABS of 50 runs by i-awGA are better than each of the other fitness assignment approach. In addition, we do not say the efficiency of the approach, only depend on the performance measure $|S_j|$ or $R_{\mathrm{NDS}}(S_j)$. We can have worst results when compared to another run with a low $R_{\mathrm{NDS}}(S_j)$. Therefore, we show the proposed i-awGA outperform another approach with the efficiency both of the performance measure $|S_j|$ or $R_{\mathrm{NDS}}(S_j)$. In Table 3, the values of $|S_j|$ are given as rational numbers, though the value of $|S_j|$ was defined as the integer number. Because we give an average of $|S_j|$ with 50 runs for comparing the different approaches. Furthermore, the values of $|S_j|$ increases (or decreases) depended on the characteristic of different testing data.

Table 1. The ABS of 50 runs by different ga parameter settings with different genetic representations

		ahnGA			priGA			
ID	Optimal	Para1	Para2	Para3	Para1	Para2	Para3	Auto-timing
1	47.93	47.93	47.93	47.93	47.93	47.93	47.93	47.93
2	210.77	232.38	234.36	244.64	224.82	234.91	228.72	224.09
3	1.75	2.69	2.71	2.83	2.68	2.73	2.79	2.64
4	17.53	37.6	39.43	47.26	36.1	35.3	34.08	34.6
5	54.93	60.77	62.26	65.35	57.26	57.42	58.5	56.87
6	234.45	276.72	288.71	295.77	269.23	268.52	273.16	270.66
7	1.83	2.4	2.66	3.31	2.01	2.27	2.32	1.98
8	22.29	47.29	49.58	57.04	41.48	45.89	44.17	41.9
9	70.97	-	-	-	72.29	75.74	77.27	70.97
10	218.78	-	-	-	276.56	276.15	284.85	272.1
11	3.82	-	-	-	5.85	6.91	6.41	5.78
12	20.63	-	-	-	60.14	57.52	61.53	52.18

"-" means out of memory error

4 Bicriteria Spanning Tree (bST) Model

Modeling and design of large communication and computer networks has always been an important area to both researchers and practitioners. The interest in developing efficient design models and optimization methods has been stimulated by high deployment and maintenance costs of networks, which make good network design potentially capable of securing considerable savings [33].

Table 2. The ABS of 50 runs by different fitness assignments

	$\lvert S_j \rvert$				$R_{\text{NDS}}(S_j)$				$Dl_R(S_j)$			
ID	spEA	nsGA	rwGA	i-awGA	spEA	nsGA	rwGA	i-awGA	spEA	nsGA	rwGA	i-awGA
1	1.64	1.7	1.64	1.84	1	1	1	1	0	0	0	0
2	5	5.08	4.98	5.64	0.18	0.16	0.22	0.38	0.18	0.23	0.17	0.1
3	3.3	3.04	3.22	3.48	0.91	0.93	0.92	0.91	0	0	0	0
4	7.36	7.4	7.12	7.46	0.04	0.02	0.04	0.04	0.06	0.06	0.05	0.05
5	3.26	3.22	3.12	3.46	1	1	1	1	0	0	0	0
6	1.74	2.4	2.2	1.54	0.28	0.14	0.18	0.3	0.17	0.24	0.22	0.15
7	4.16	3.96	3.66	3.7	0.52	0.59	0.66	0.68	0.4	0.42	0.4	0.05
8	5.9	4.8	5.3	5.16	0.05	0.13	0.07	0.1	1.1	0.89	0.96	0.86
9	1.16	1.24	1.28	1.36	0.99	0.96	0.91	0.99	0	0.01	0.01	0
10	2.6	2.42	2.62	2.3	0.11	0.18	0.16	0.33	1.17	0.76	0.99	0.59
11	2.86	2.9	2.7	3.22	0.31	0.3	0.3	0.43	0.01	0.04	0.01	0
12	5.82	6.02	6.14	6.2	0.03	0.03	0.04	0.05	0.19	0.19	0.2	0.19

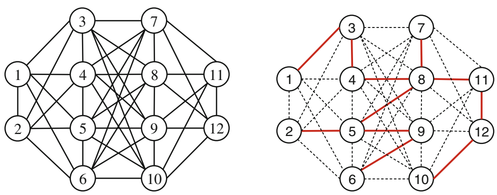
The Minimum Spanning Tree (MST) problem is one of the best-known network optimization problems which attempt to find a minimum cost tree network that connects all the nodes in the communication network. The links or edges have associated costs that could be based on their distance, capacity, quality of line, etc.

In the real world, the MST is often required to satisfy some additional constrain for designing communication networks such as the capacity constraints on any edge or node, degree constraints on nodes, and type of services available on the edge or node. This additional constraint often makes the problem NP-hard. In addition, there are usually such cases that one has to consider simultaneously multicriteria in determining a MST, because there are multiple attributes defined on each edge, has become subject to considerable attention. Almost every important real-world decision making problem involves multiple and conflicting objectives [17].

In this paper, we are considering a bicriteria spanning tree (bST) model. The bST is to find a set of links with the two conflicting objectives of minimizing communication cost and minimizing the transfer delay and the constraint of network capacity is met. This problem can be formulated as the multiobjective capacitated minimum spanning tree (mcMST) problem, and is a NP-hard.

4.1 Mathematical Formulation

The communication network is modeled using an edge-weighted undirected graph $G = (V, E, C, D, W)$ with $n = |V|$ nodes and $m = |E|$ edges, $c_{ij} \epsilon C = [c_{ij}]$, $d_{ij} \epsilon D = [d_{ij}]$ and $w_{ij} \epsilon W = [w_{ij}]$ represent the cost, delay and weight of each edge $(i, j) \epsilon E$, where the variable is restricted to be a nonnegative real number, respectively. And u_i is given weight capacity of each node i. Figure 3 presents a simple network with 12 nodes and 40 arcs.



(a) a simple network (b) a spanning tree

Fig. 3. A simple network with 12 nodes and 40 edges

The decision variables in the bicriteria communication spanning tree problem (bSTP) are the minimum cost z_1 with minimum delay z_2. Mathematically, the problem is reformulated as a constrained Minimal Spanning Tree (MST) problem is as follows:

$$\min\ z_1(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} \tag{5}$$

$$\min\ z_2(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}x_{ij} \tag{6}$$

$$\text{s. t. } \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} = n-1 \tag{7}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} \leq |S| - 1 \text{ for any set } S \text{ of nodes} \tag{8}$$

$$\sum_{j=1}^{n} w_{ij}x_{ij} \leq u_i, \forall i \tag{9}$$

$$x_{ij} \in \{0,1\}, \forall i,j. \tag{10}$$

In this formulation, the 0-1 variable x_{ij} indicates whether we select edge (i,j) as part of the chosen spanning tree (note that the second set of constraints with $|S| = 2$ implies that each $x_{ij} \leq 1$). The constraint (7) is a cardinality constraint implying that we choose exactly $n-1$ edges, and the packing constraint (8) implies that the set of chosen edges contain no cycles (if the chosen solution contained a cycle, and S were the set of nodes on a chosen cycle, the solution would violate this constraint). The constraint (9) guarantees that the total link weight of each node i does not exceed the upper limit W_i.

4.2 GA Approaches for bST Model

Recently, GA and other Evolutionary Algorithms (EAs) have been used extensively in a wide variety of communication network design problems. For example, Mathur [32] Voβ, Stefan et al. [46] gave the EAs for the capacitated MST problem occurring in telecommunication applications, Ruiz and Duhamel [10,37] investigated the different encoding methods and gave the performance analyzes for the degree-constraint MST problem in communication network, respectively. Liang [29], Torkestani et al. [43], Rezaei et al. [45] investigated the multicriteria communication design problems with GA approaches and other EAs. GAs were also applied to solve other communication network design problem such as two graph problem, one-max tree problem and communication spanning tree problem [44].

We need to consider these critical issues carefully when designing an appropriate encoding method so as to build an effective GA. How to encode a spanning tree T in a graph G is critical for developing a GA to network design problems, it is not easy to find out a nature representation. We summarized the several kinds of classification of encoding methods as follows:

Characteristic Vectors-based Encoding: Davis et al. [5], Piggott and Suraweera [34] have used binary-based encoding method to represent spanning trees in GAs.

Edge-based Encoding: Edge-based encoding is an intuitive representation of a tree. General, an edge-based encoding requires space proportional to n and the time complexities is $O(m)$. Edge-based encoding and binary-based encoding have very similar performance in theory. However, there are $2n(n-1)/2$ possible values for a tree, and only a tiny fraction of these chromosomes represents feasible solutions, and weaken the encoding heritability.

Prefer Number-based Encoding: Many researchers have encoded spanning trees as Prüfer numbers in GAs, called Prüfer number-based encoding for a variety of problems. These include the degree-constrained MST problems, leaf-constrained MST problem, and multicriteria MST problems, etc.

Predecessor-based Encoding: Abuali et al. [35] proposed a predecessor-based encoding that applies to spanning trees. It is a more compact representation of spanning trees by the predecessor or determinant encoding, in which an arbitrary node in G is designated the root, and a chromosome lists each other node's predecessor in the path from the node to the root in the represented spanning tree: if pred(i) is j, then node j is adjacent to node i and nearer the root. Thus, a chromosome is string of length $n-1$ over $1, 2, \cdots, n$, and when such a chromosome decodes a spanning tree, its edges can be made explicit in time that is O(nlogn).

PrimPred-based Encoding: We improved the predecessor-based encoding that adopted Prim's algorithm in chromosome generating procedure. Prim's algorithm implements the greedy-choice strategy for minimum spanning tree. Starting with an empty tree (one with no edges), the algorithm repeatedly adds the lowest-weight edge (u, v) in G such that either u or v, but not both, is already connected to the tree. Considering the characteristic of predecessor-based encoding, they proposed a new crossover and mutation operators. These operators offer locality, heritability, and computational efficiency. An example of generated PrimPred-based chromosome and its decoded spanning tree are shown in Fig. 4 for the undirected network shown in Fig. 3.

node id k:	2	3	4	5	6	7	8	9	10	11	12
chromosome $v(k)$:	5	1	3	8	9	8	4	5	12	8	11

$T = \{(1, 3), (2, 5), (3, 4), (4, 8), (5, 8), (5, 9), (6, 9), (7, 8), (8, 11), (10, 12), (11, 12)\}$

Fig. 4. An example of the Prim Pred-based chromosome

4.3 Experiments and Discussion

In this section, first, the proposed PrimPred-based GA is compared with Prefer Number-based Encoding, Edge-based Encoding for solving several large scale minimum spanning tree problems. Afterwards, compare the different fitness assignment approaches by multiobjective GA for solving bSTP.

(1) Different Genetic Representations
For examining the effectiveness of different encoding methods, we applied Prefer Number-based encoding method and Edge-based encoding method on 6 test problems [30]. We are combining the Prefer Number-based encoding with one-cut point crossover and swap mutation, and combining the Edge-based encoding using two kinds of mutation operators which included in [48], and for initializing the chromosomes based on the edge set, we combining the Raidl and Julstrom's PrimRST (Prim Random Spanning Tree). Each algorithm was run 20 times using different initial seeds for each test problems. The GA parameter is setting as follows:

Population size: popSize = 10;
Crossover probability: $p_C = 0.30, 0.50$ or 0.70;
Mutation probability: $p_M = 0.30, 0.50$ or 0.70;
Maximum generation: maxGen = 1000;

The experimental study was realized to investigate the effectiveness of the different encoding method; the interaction of the encoding with the crossover operators and mutation operators; and the parameter settings affect its performance. Table 3 gives computational results for four different encoding methods on six test problems by tree kinds of parameter settings, where, edge-based 1 and edge-based 2 were considered with different mutation operators of Raidl and Julstrom's approach [35].

When we compare with columns of the best cost of four encoding methods, it is possible to see that whereas the Prefer number-based approach is the fastest than others, it is difficult to consist mostly of substructures of their parents' phenotypes (poor heritability), and the results is very far from the best one. The two kinds of mutation is used in edge-based encoding, the second one (depends on the cost) gives better performance than the first one. For considering the computational cost (CPU time), because of the LowestCost mutation in proposed approach, spend the most CPU times to find the edge with lowest cost, it is always longer than other algorithms. However, our algorithm developed in this study gives best cost than other algorithms.

(2) Different Fitness Assignment Approaches
The second experimental study, we show performance comparisons of multiobjective GAs for solving bSTP by different fitness assignment approaches, there are spEA, nsGAII, rwGA and i-awGA. The data in test problem was generated randomly. In each GA approach, PrimPred-based encoding was used, and Prim-based crossover and LowestCost mutation were used as genetic operators. GA parameter settings were taken as follows:

Population size: popSize = 20;
Crossover probability: $p_C = 0.70$;
Mutation probability: $p_M = 0.50$;
Stopping criteria: evaluation of 5000 solutions.

We compare i-awGA with spEA, nsGAII and rwGA trough computational experiments on the 40-node/1560-arc test problem under the same stopping condition (i.e., evaluation of 5000 solutions). Each algorithm was applied to

Table 3. Performance comparisons with different GA approaches.

Test problem	Optimal solutions					Prüfernum-based		Edge-based1		Edge-based2		PremPred-based	
		n	m	p_C	p_M	avg	CPU-t	avg	CPU-t	avg	CPU-t	avg	CPU-t
1	470	40	780	0.3	0.3	1622.2	72.2	1491.8	1075.2	495.6	1081.4	470	1100.2
				0.5	0.5	1624.4	87.6	1355.8	2184.4	505.8	2175	470	2256.4
				0.7	0.7	1652.6	134.8	1255.2	3287.4	497.6	3281.4	470	3316
2	450	40	780	0.3	0.3	1536.6	74.8	1458.2	1118.6	471.6	1093.8	450	1160.2
				0.5	0.5	1549.2	78.2	1311.4	2190.8	480.2	2175	450	2200.2
				0.7	0.7	1564.4	122	1184.4	3287.6	466.4	3262.4	450	3275
3	820	80	3160	0.3		0.3 3880.4	150	3760.2	5037.8	923.2	5059.6	820	5072
				0.5	0.5	3830	184.4	3692	10381.2	871	10494.2	820	10440.6
				0.7	0.7	3858.2	231.2	3483.8	16034.8	899.2	15871.8	820	15984.6
4	802	80	3160	0.3		0.3 3900.6	131.4	3853	5125	894.6	4934.2	802	5071.8
				0.5	0.5	3849.6	206.2	3515.2	10325.2	863	10268.8	802	10365.6
				0.7	0.7	3818.4	222	3287.2	16003	868	15965.4	802	15947.2
5	712	120	7140	0.3		0.3 5819.4	187.4	5536.6	15372	871.8	15306.4	712	15790.4
				0.5	0.5	5717.2	293.8	5141	31324.8	805.4	30781.4	712	31503.2
				0.7	0.7	5801.4	316	5035.2	47519	804.2	47047.2	712	47865.8
6	793	160	12720	0.3		0.3 7434.8	284.4	7050.4	41993.6	1353.6	42418.6	809.6	42628.2
				0.5	0.5	7361	421.8	7111.6	87118.8	1061.6	86987.4	793	86828.4
				0.7	0.7	7517	403.2	6735	163025	955.4	161862.4	793	154731.2

Table foot note (with superscript)

Table 4. Performance evaluation of fitness assignment approaches for the 40-node/1560-arc test problem.

# of eval.solut.	$\|S_j\|$				$R_{\mathrm{NDS}}(S_j)$				$Dl_R(Sj)$			
	spEA	nsGA II	rwGA	i-awGA	spEA	nsGA II	rwGA	i-awGA	spEA	nsGA II	rwGA	i-awGA
50	31.45	30.4	32.6	36.2	0.34	0.31	0.36	0.39	178.85	200.47	182.03	162.57
500	42.4	45.6	43.2	47.6	0.42	0.45	0.4	0.52	162.97	151.62	160.88	157.93
2000	46.6	52.2	45.3	55.5	0.54	0.61	0.58	0.66	118.49	114.6	139.4	92.41
5000	51.2	54.4	50.3	60.7	0.64	0.7	0.62	0.73	82.7	87.65	117.48	77.98

each test problem 10 times and gives the average results of the 3 performance measures (i.e., the number of obtained solutions $|S_j|$, the ratio of nondominated solutions $R_{\mathrm{NDS}}(S_j)$, and the average distance D1R measure). In Table 4, better results of all performance measures were obtained from the i-awGA than other fitness assignment approach.

5 Conclusion

In this paper, we investigated with a broad spectrum of bicriteria network optimization models, analyze the recent related researches, design and validate effective EA approaches for the typical network models: bicriteria shortest path (bSP) model, bicriteria spanning tree (bST) model, bicriteria network flow (bNF) model. Focus on the broad applications of proposed evolutionary algorithms (EAs) to network design with large-scale optimization.

For bSP model, we introduced an efficient genetic representation using the priority-based encoding method; based on the characteristic of this representation, we propose a new crossover operator, called weight mapping crossover (WMX), adopt insertion mutation and immigration operator. We gave effective analysis of different evolutionary approaches for bSP model dedicated to calculate nondominated paths for the minimum total cost and the minimum transmission delay. For bST model, we investigated different GA approaches for solving minimum spanning tree (MST) problems, and introduced a new genetic representation using the PrimPred-based encoding method. And we also introduced a local search technique into GA loops, called a LowestCost mutation operator. For bNF model, we introduced a new multiobjective genetic algorithm (moGA) to solve the problem with two conflicting objectives to minimize the total cost and maximize the total flow simultaneously. We combine the priority-based encoding method with a special decoding to overcome the special difficulties of bNF model.

Acknowledgements. This work is partly supported by the National Natural Science Foundation of China under Grant 61572100, and in part by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS) No. 15K00357.

References

1. Abbasi S, Taghipour M (2015) An ant colony algorithm for solving bi-criteria network flow problems in dynamic networks. IT Eng 3(5):34–48
2. Ahuj RK, Magnanti TL, Orlin JB (1993) Network flows. Prentice Hall, New Jersey
3. Climaco JCN, Craveirinha JMF, Pascoal MMB (2003) A bicriterion approach for routing problems in multimedia networks. Networks 41(4):206–220
4. Craveirinha J, Maco J et al (2013) A bi-criteria minimum spanning tree routing model for mpls/overlay networks. Telecommun Syst 52(1):1–13
5. Davis L, Orvosh D et al (1993) A genetic algorithm for survivable network design. In: Proceedings of 5th international conference on genetic algorithms, pp 408–415
6. Deb BK (2010) Optimization for engineering design: algorithms and examples. Prentice-Hall, New Delhi
7. Deb K (1989) Genetic algorithms in multimodal function optimization. Master's thesis, University of Alabama
8. Deb K, Thiele L et al (2001) Scalable test problems for evolutionary multiobjective optimization. Wiley, Chichester
9. Dorigo M (1992) Optimization, learning and natural algorithms. PhD thesis, Politecnico di Milano
10. Duhamel C, Gouveia L et al (2012) Models and heuristics for the k-degree constrained minimum spanning tree problem with node-degree costs. Networks 60(1):1–18
11. Fawcett H (2014) Manual of political economy. Macmillan and co., London
12. Fogel LJ, Owens AJ, Walsh MJ (1966) Artificial intelligence through simulated evolution. Wiley, New York
13. Fonseca C, Fleming P (1995) An overview of evolutionary algorithms in multiobjective optimization. IEEE Trans Evol Comput 3(1):1–16

14. Garey MR, Johnson DS (1979) Computers and intractability: a guide to the theory of np-completeness. W.H. Freeman, New York
15. Gen M (2006) Genetic algorithms and their applications. Springer, London
16. Gen M, Cheng R, Oren SS (2000) Network design techniques using adapted genetic algorithms. Springer, London
17. Gen M, Cheng R, Lin L (2008) Network models and optimization: multiobjective genetic algorithm approach. Springer, London
18. Goldberg DE (1989) Genetic algorithms in search, optimization, and machine learning. Addison-Wesley Pub. Co., Boston
19. Hansen P (1979) Bicriterion path problems. In: Proceedings of 3rd conference multiple criteria decision making theory and application, pp 109–127
20. Hao X, Gen M et al (2015) Effective multiobjective EDA for bi-criteria stochastic job-shop scheduling problem. J Intell Manufact 28:1–13
21. Holland J (1975) Adaptation in Natural and Artificial System. MIT Press, Ann Arbor
22. Holland JH (1976) Adaptation. In: Rosen R, Snell FM (eds) Progress in theoretical biology IV
23. Hwang CL, Yoon K (1994) Multiple attribute decision making. Springer, Heidelberg
24. Ishibuchi H, Murata T (1998) A multi-objective genetic local search algorithm and its application to flowshop scheduling. Comput Ind Eng 28(3):392–403
25. Kennedy J, Eberhart R (2011) Particle swarm optimization. In: Proceeding of the IEEE international conference on neural networks, Piscataway, pp 1942–1948
26. Kennedy J, Eberhart R (2011) Particle swarm optimization. Morgan Kaufmann, San Francisco
27. Koza JR (1992) Genetic programming, the next generation. MIT Press, Cambridge
28. Koza JR (1994) Genetic programming II (videotape): the next generation. MIT Press, Cambridge
29. Liang W, Schweitzer P, Xu Z (2013) Approximation algorithms for capacitated minimum forest problems in wireless sensor networks with a mobile sink. IEEE Trans Comput 62(10):1932–1944
30. Lin L (2006) Node-based genetic algorithm for communication spanning tree problem. IEICE Trans Commun 89(4):1091–1098
31. Lin L, Gen M (2008) An effective evolutionary approach for bicriteria shortest path routing problems. IEEJ Trans Electron Inf Syst 128(3):416–423
32. Mathur R, Khan I, Choudhary V (2013) Genetic algorithm for dynamic capacitated minimum spanning tree. Comput Technol Appl 4(3):404
33. Medhi D, Pioro M (2004) Routing, flow, and capacity design in communication and computer networks. Morgan Kaufmann Publishers, San Francisco
34. Piggott P, Suraweera F (1995) Encoding graphs for genetic algorithms: an investigation using the minimum spanning tree problem. Springer, Heidelberg
35. Raidl GR, Julstrom B (2003) Edge-sets: an effective evolutionary coding of spanning trees. IEEE Trans Evol. Comput. 7(3):225–239
36. Rechenberg I (1973) Optimieriung technischer Systeme nach Prinzipien der biologischen Evolution. Frommann-Holzboog, Stuttgart
37. Ruiz E, Albareda-Sambola M et al (2015) A biased random-key genetic algorithm for the capacitated minimum spanning tree problem. Comput Oper Res 57:95–108
38. Schaffer JD (1985) Multiple objective optimization with vector evaluated genetic algorithms. In: International conference on genetic algorithms, pp 93–100
39. Schwefel HPP (1995) Evolution and Optimum Seeking: The Sixth Generation. Wiley, New York

40. Srinivas N, Deb K (1995) Multiobjective function optimization using nondominated sorting genetic algorithms. IEEE Trans Evol Comput 2(3):221–248
41. Steiner S, Radzik T (2008) Computing all efficient solutions of the biobjective minimum spanning tree problem. Comput Oper Res 35(1):198–211
42. Steuer RE (1986) Multiple criteria optimization theory computation and application Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Wiley, New York
43. Torkestani JA (2013) Degree constrained minimum spanning tree problem: a learning automata approach. J Supercomput 64(1):226–249
44. Torkestani JA, Meybodi MR (2012) A learning automata-based heuristic algorithm for solving the minimum spanning tree problem in stochastic graphs. J Supercomput 59(2):1035–1054
45. Torkestani JA, Rezaei Z (2013) A learning automata based approach to the bounded-diameter minimum spanning tree problem. J Chin Inst Eng 36(6):749–759
46. Voβ S, Martello S, et al (2012) Meta-heuristics: Advances and trends in local search paradigms for optimization
47. Zandieh M (2011) A novel imperialist competitive algorithm for bi-criteria scheduling of the assembly flowshop problem. Int J Prod Res 49(11):3087–3103
48. Zhang Q, Ding L (2016) A new crossover mechanism for genetic algorithms with variable-length chromosomes for path optimization problems. Expert Syst Appl 60:183–189
49. Zitzler E, Laumanns M, Thiele L (2001) SPEA2: improving the strength pareto evolutionary algorithm for multiobjective optimization. In: Evolutionary methods for design, optimization and control with applications to industrial problems. Proceedings of the Eurogen 2001, Athens, Greece, September

Study of County Level Government's E-Government Efficiency Evaluation in Sichuan Province Based on DEA

Yu Liu(✉), Jiawei Liu, Zhaohui Zhang, and Liangqing Zhang

School of Political Science and Public Administration,
University of Electronic Science and Technology of China,
Chengdu 611731, People's Republic of China
laujames_work@163.com

Abstract. The county government plays a fundamental role in the Chinese governance. The e-government system of it can directly reflect the quality and efficiency of its service. Thus, it's significant to evaluate the efficiency of the e-government at county level. Nowadays, there're plenty of methods to do this kind of evaluation. This paper mainly uses principal component analysis to establish the index system, and adopts the CCR and BCC models, two of the data envelopment analysis DEA, to reflect the current input and output efficiency of 147 counties' e-government in Sichuan Province. The analysis reveals that most of the e-governments' comprehensive efficiency, technical efficiency and scale efficiency are non-DEA effective in county level governments of Sichuan Province, and the differences in development among them are quite distinctive. What's more, the local governments' information department should embark on improving their investment scale and technical efficiency in order to enhance the efficiency of e-government.

Keywords: E-government · County government · Data envelopment analysis · Efficiency

1 Introduction

The county government has long been the basis of administration management and state governance in China. It faces to the grass-roots and serves them directly. Thus, the efficiency of its operation can greatly influence the overall quality of people's life and production. Nowadays, as an effective way of improving administrative efficiency and transparency, the e-government has been praised highly by many countries. And China has also launched the Government Online Project since 1999. In fact, the construction of e-government within the county level is a very important part of China's e-government development, as well as the terminal node to the provincial level e-government, and governments should pay great attention to e-government construction at county level for constructing mature e-government system structure [7]. To promote the healthy and

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_32

sustainable development of a county e-government, an evaluation of its efficiency which may help improve its service quality is strongly required.

Evaluation systems like the Global E-government Evaluation from the UN [11], the "eEurope" Strategic Assessment System from the EU [3] and the Overall e-government Maturity Assessment from the Accenture company [1] are international representatives. While in China, it also has assessment models like the CCW Research [12], the China Software Testing Center [4], the CCID Consulting [5] and the CCTC [2].

According to the actual situation, the E-government efficiency evaluation index system was constructed from different perspectives and the corresponding analysis methods are used to evaluate the efficiency of the E-government. Extraordinary representatives: Luo and Shi [10] introduced genetic algorithm optimized to BP neural network weights and thresholds after reducing the established evaluation index system by rough set, and established e-government website evaluation model based on genetic neural network algorithm; Rorissa, Demissie and Pardo [13] used benchmarking to assess the strengths and limitations of six frameworks for computing e-government indexes; Luna et al. [9] used data envelopment analysis (DEA) to show how efficient are Mexico state governments in their use of certain inputs to produce high quality e-government portals; Zdjelar [17] measured efficiency of Croatian regional government implementation by Balanced Scorecard Method; Song and Guan [14] utilized a super efficiency slack-based measure (SBM) model to evaluate the e-government performance of environmental protection administrations in the 16 cities of Anhui Province; Kao [8] evaluated the e-government of 21 European countries based on the concept of Pareto optimality and developed a compromise programming model to rank these nations; Wu and Guo [15] used method of DEA to evaluate e-government performances of 31 provincial government websites in China and concluded that most of these provincial government websites operated at an inefficient level and in a bad manner.

To sum up, the achievements of researches on the e-government's efficiency provide a useful reference to the study. However, the evaluation of Chinese e-government's efficiency has not yet had a uniform index system. Most researches are focusing on the overall efficiency of the provincial research objects and few investigate the input-output efficiency of e-government from the perspective of e-government at county level. Therefore, this paper attempts to use principal component analysis to build the evaluation index system of county level e-government's efficiency, and adopt the DEA model to evaluate the efficiency of e-government in 147 counties (including county level cities) of Sichuan Province to provide decision-making reference for the promotion of county level e-government's efficiency.

The remainder of this paper is organized as follows. In the first half of Sect. 2, we briefly discuss the principle of principal component analysis and the practical application value of it. In the second half of Sect. 2, we narrate the construction of CCR model and BCC model of DEA which will be used for e-government efficiency evaluation. The process of e-government efficiency evaluation is presented in Sect. 3 which includes the samples selection, data collection, the establishment of evaluation index system and the model solution. On the basis of the solution, the analysis of the current situation in Sichuan Province and the suggestions for the ineffective county level government are given in Sect. 4. Finally, we conclude with a summary of this paper and prospect about evaluation of e-government in Sect. 5.

2 The Method of E-Government Efficiency Evaluation

2.1 Principal Component Analysis

On the premise of the least information loss, principal component analysis is a statistical procedure that uses the thought of dimension reduction to convert a set of multiple indexes with possible correlation into a few linearly uncorrelated comprehensive indexes.

The basic principles of principal component analysis are as follows:

We assume that there are p objects that each object is described by n indexes. In this way, we can constitute a data matrix:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{pmatrix} = (X_1, X_2, \cdots, X_n). \tag{1}$$

$X_1, X_2, \cdots, X_n$ are defined as original variable indexes and $F_1, F_2, \cdots, F_m (m \le n)$ are defined as new variable indexes. By using the n vectors $X_1, X_2, \cdots, X_n$ of data matrix to make a linear combination, then we can get:

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n \\ F_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2n}X_n \\ \qquad\qquad \cdots \qquad \cdots \\ F_m = a_{m1}X_1 + a_{m2}X_2 + \cdots + a_{mn}X_n \end{cases} \tag{2}$$

Coefficient a_{ij} meets:

- F_i is linearly independent with $F_j (i \neq j; i, j = 1, 2, \cdots, n)$;
- F_1 has the greatest variance of the linear combination of $X_1, X_2, \cdots, X_n$ mentioned above. F_2 has the greatest variance of the linear combination of $X_1, X_2, \cdots, X_n$ which are independent with F_1. The rest could be deduced in this way. F_m has the greatest variance of the linear combination of $X_1, X_2, \cdots, X_n$ which are independent with $F_1, F_2, \cdots, F_{m-1} (m \le n)$.

New variable indexes $F_1, F_2, \cdots, F_m (m \le n)$ are called the first, second, $\cdots$, m^{th} principal component respectively.

2.2 Data Envelopment Analysis

Based on the relative efficiency concept, the data envelopment analysis is a non-parametric statistical method employing convex analysis and linear programming as tools to calculate the relative efficiency between the same type of multi-input and multi-output decision making units. The relative efficiency is used to evaluate whether the decision-making unit is effective.

(1) CCR Mode
Assume that there are n decision-making units (hereinafter referred to as DMU), each DMU has n inputs and m outputs. The input vector of DMU_j is $X_j = (x_{1j}, x_{2j}, \cdots, x_{mj})^T \geq 0, X_j \in R^+$ while output vector is $Y_j = (y_{1j}, y_{2j}, \cdots, y_{sj})^T \geq 0, Y_j \in R^+$ Building a model with a non-Archimedes infinitely small to evaluate the relative efficiency of the DMU_j:

$$\min\left[\,\theta - \varepsilon(e_1^T s^- + e_2^T s^+)\right]$$
$$s.t.\ \begin{cases} \sum\limits_{j=1}^{n} \lambda_j x_j + s^- = \theta x_0 \\ \sum\limits_{j=1}^{n} \lambda_j y_j - s^+ = y_0 \\ \lambda_j \geq 0, j = 1, 2, \cdots, n \\ s^+ \geq 0, s^- \geq 0 \end{cases} \tag{3}$$

ε is non-Archimedes infinitely small. s^- is the slack variable while s^+ is the remaining variable. Assuming that Eq. (3) has the best solution $\theta_0, \lambda_0, s_0^-, s_0^+$ for DMU_{j0}. If $\theta_0 < 1$, DMU_{j0} will be non-DEA effective. If $\theta_0 = 1$ and s_0^- or $s_0^+ \neq 0$, DMU_{j0} will be weakly DEA effective. If $\theta_0 = 1$ and DMU_{j0} will be DEA effective.

(2) BCC Model
BCC model is the extension of fixed scale benefit data envelopment analysis. Adding the convexity condition $\sum_{j=1}^{n} \lambda_j = 1$ to the linear programming under fixed scale benefit model (CCR model) is the BCC model:

$$\min\left[\,\theta - \varepsilon(e_1^T s^- + e_2^T s^+)\right]$$
$$s.t.\ \begin{cases} \sum\limits_{j=1}^{n} \lambda_j x_j + s^- = \theta x_0 \\ \sum\limits_{j=1}^{n} \lambda_j y_j - s^+ = y_0 \\ \sum\limits_{j=1}^{n} \lambda_j = 1 \\ \lambda_j \geq 0, j = 1, 2, \cdots, n \\ s^+ \geq 0, s^- \geq 0. \end{cases} \tag{4}$$

3 The Process of E-Government Efficiency Evaluation

3.1 Samples, Indexes and Data Collection

The decision-making units in DEA should have the same type of characteristics. In order to evaluate the relative effectiveness of e-government at the county level in Sichuan Province, this paper selects the 147 districts and counties (including county level cities) in Sichuan Province as the objects of the research.

In accordance with the requirements of the DEA method, we need to establish an index system of input and output based on the comprehensive consideration of region informatization level, economic development, cultural and educational level and other factors. There are nine categories of input indexes in this paper which specifically include fixed telephone subscribers, the number of mobile phone users, the value of the tertiary industry income, GDP, total fiscal revenue, governmental investment in science and education, the proportion of urban population, the number of students in colleges and universities, the number of colleges and universities. These index data are mainly obtained from "Sichuan Province Statistical Yearbook" (2014) [16] and "The Statistics Bulletin of the National Economy and Social Development" (2013) issued by the various districts and counties (including county level cities).

To select proper output indicators, we referenced the objective indexes derived from CCID consulting assessment of government website performance and The United Nations for the global objective of e-government assessment. This study revolves around the authoritative evaluation results to screen the indexes. The five main evaluation indexes consist of government information publicity, work services, public participation, the website management, application of new technology. Output index data are mainly obtained from "The total Report of the Sichuan Government Website Performance Evaluation" (2013) [6].

3.2 Principal Component Analysis

The principal component analysis of 9 input indexes and 5 output indexes data were conducted by SPSS software (Tables 1 and 2).

According to extracts of the principal component to calculate the principal component load, the formula is:

$$l_{ij} = p(z_i, x_j) = \sqrt{\lambda_i} a_{ij} (i = 1, 2, \cdots, k, k \text{ as main composition}, \quad j = 1, 2, \cdots, 147) \tag{5}$$

Then we can get the principal component index data of inputs and outputs.

Table 1. The result of PCA of the input indexes

Principal component analysis	Eigenvalue
$F1$	5.963

Table 2. The result of PCA of the output indexes

Principal component analysis	Eigenvalue
E1	2.469
E2	1.004

Table 3. The results of e-government input and output efficiency evaluation (1)

City or Autonomous prefecture	County	CCR	BCC	Scale efficiency
Chengdu City	Jinjiang District	0.919	1	0.919
	Qingyang District	0.581	0.581	0.999
	Jinniu District	0.557	0.582	0.957
	Wuhou District	1	1	1
	Chenghua District	0.832	0.995	0.836
	Longquanyi District	0.627	0.627	0.999
	Qingbaijiang District	0.226	0.226	1
	Xindu District	0.583	0.868	0.672
	Wenjiang District	0.464	0.465	0.999
	Dujiangyan City	0.318	0.318	0.998
	Pengzhou City	0.261	0.262	0.999
	Qionglai City	0.17	0.17	0.999
	Chongzhou City	0.193	0.385	0.5
	Jintang County	0.18	0.18	0.999
	Shuangliu County	0.673	0.673	1
	Pixian County	0.587	1	0.587
	Dayi County	0.185	0.187	0.991
	Pujiang County	0.226	0.226	0.999
	Xinjin County	0.318	0.77	0.413
Guangan City	Guangan District	0.201	0.202	0.998
	Linshui County	0.179	0.18	0.995
	Wusheng County	0.157	0.158	0.996
	Yuechi County	0.158	1	0.158
	Huayun City	0.139	0.14	0.997
Bazhong City	Bazhou District	0.185	0.185	1
	Pingchang County	0.153	0.153	0.996
	Tongjiang County	0.144	0.144	0.997
	Nanjiang County	0.175	0.175	0.998
Meishan City	Dongpo District	0.257	0.258	0.999
	Pengshan District	0.138	0.218	0.633
	Renshou County	0.221	0.221	0.998
	Qingshen County	0.088	0.113	0.777
	Hongya County	0.092	0.1	0.922

(*Continued*)

Table 3. (*Continued*)

City or Autonomous prefecture	County	CCR	BCC	Scale efficiency
Tibetan Autonomous Prefecture of Garze	Kangding County	0.199	0.212	0.936
	Luding County	0.146	0.213	0.686
	Jiulong County	0.117	0.218	0.537
	Dawu County	0.088	0.159	0.551
	Litang County	0.095	0.152	0.622
	Daocheng County	0.132	0.26	0.507
	Yajiang County	0.104	0.204	0.511
Liangshan Autonomous Prefecture	Xichang City	0.341	0.342	0.998
	Dechang County	0.116	0.146	0.793
	Huili County	0.162	0.162	0.999
	Huidong County	0.12	0.137	0.876
	Ningnan County	0.123	0.151	0.817
	Jinyang County	0.077	0.145	0.529
	Butuo County	0.083	0.16	0.516
	Zhaojue County	0.096	0.159	0.599
	Leibo County	0.124	0.18	0.689
	Ganluo County	0.105	0.181	0.582
	Yuexi County	0.098	0.148	0.666

On account of that the result of the principal component analysis is negative number, and DEA model requires that the data is not negative, so the results of the principal component calculated by poor transformation normalized to [1, 10] interval. The range transform formula is:

$$n_{\text{new}} = 1 + 9 \times [(n - n_{\min})/(n_{\max} - n_{\min})], \tag{6}$$

where $n_{\max}$ and $n_{\min}$ represent the maximum and minimum values of the column data respectively (Tables 3 and 4).

3.3 The Solution of the DEA Model

According to the principal component analysis method, the main components of input and output index data are obtained. After that, we use the DEAP 2.1 to solve 147 counties' (including districts' and cities') input-output efficiency of e-government with CCR and BCC model, and we obtain the comprehensive efficiency, pure technical efficiency and scale efficiency of all counties' (including districts' and cities') e-government. All results are shown in Table 5.

Table 4. The results of e-government input and output efficiency evaluation (2)

City or Autonomous prefecture	County	*CCR*	*BCC*	Scale efficiency
Yibin City	Cuiping District	0.386	0.387	0.999
	Nanxi District	0.166	0.19	0.874
	Yibin County	0.184	0.185	0.995
	Jiang'an County	0.153	0.154	0.995
	Gongxian County	1	1	1
	Junlian County	0.151	0.168	0.903
	Xingwen County	0.159	0.164	0.967
	Pingshan County	0.106	0.161	0.659
Ya'anCity	Yucheng District	0.423	0.424	0.998
	Hanyuan County	0.171	0.145	0.848
	Shimian County	0.117	0.151	0.774
	Lushan County	0.114	0.152	0.747
Nanchong City	Shunqing District	0.346	0.347	0.999
	Gaoping District	0.134	0.134	0.998
	Xichong County	0.124	0.124	0.996
	Nanbu County	0.16	0.161	0.998
	Yilong County	0.141	0.142	0.998
	Yingshan County	0.147	0.148	0.995
	Peng'an County	0.107	0.108	0.998
Luzhou City	Jiangyang District	0.282	0.283	0.998
	Naxi District	0.261	0.262	0.999
	Luxian County	0.216	0.217	0.998
	Gulin County	0.137	0.137	0.998
Dazhou City	Tongchuan District	0.277	0.278	0.998
	Dachuan District	0.234	0.234	1
	Xuanhan County	0.267	0.267	1
	Kaijiang County	0.166	0.171	0.969
	Dazhu County	0.276	0.277	0.996
	Quxian County	0.25	0.251	0.997
	Wanyuan City	0.196	0.409	0.48
Aba Tibetan and Qiang Autonomous Prefecture	Maerkang County	0.171	0.189	0.903
	Jinchuan County	0.089	0.151	0.594
	Xiaojin County	0.152	0.181	0.841
	Aba County	0.088	0.161	0.55
	Ruoergai County	0.101	0.132	0.771

(*Continued*)

Table 4. (*Continued*)

City or Autonomous prefecture	County	*CCR*	*BCC*	Scale efficiency
	Hongyuan County	0.128	0.218	0.587
	Rangtang County	0.085	0.175	0.485
	Wenchuan County	0.124	0.145	0.858
	Lixian County	0.096	0.159	0.599
	Maoxian County	0.108	0.147	0.735
	Songpan County	0.114	0.177	0.645
	Jiuzhaigou County	0.103	0.133	0.775
	Heishui County	0.095	0.163	0.585
Guangyuan City	Lizhou District	0.248	0.249	1
	Zhaohua District	0.128	0.271	0.473
	Chaotian District	0.089	0.115	0.777
	Wangcang Distric	0.18	0.183	0.98
	Qingchuan County	0.17	0.232	0.732
	Jiange County	0.19	0.196	0.967
	Cangxi County	0.191	0.192	0.996

Table 5. The results of e-government input and output efficiency evaluation (3)

City or Autonomous Prefecture	County	CCR	BCC	Scale efficiency
Suining City	Chuanshan District	0.26	0.26	0.997
	Anju District	0.127	0.146	0.871
	Shehong County	0.181	0.209	0.867
	Pengxi County	0.109	0.114	0.957
	Daying County	0.121	0.134	0.905
Neijiang City	Central District	0.353	0.58	0.609
	Dongxing District	0.27	0.319	0.846
	Zizhong County	0.216	0.367	0.588
	Longchang County	0.245	0.302	0.81
	Weiyuan County	0.211	0.407	0.518
Leshan City	Central District	0.399	0.4	0.998
	Shawan District	0.18	0.191	0.943
	Wutongqiao District	0.154	0.154	0.998
	Jinkouhe District	0.102	0.152	0.67
	Qianwei County	0.186	0.187	0.997

(*Continued*)

Table 5. (*Continued*)

City or Autonomous Prefecture	County	CCR	BCC	Scale efficiency
	Jingyan County	0.122	0.145	0.839
	Jiajiang County	0.153	0.153	0.997
	Muchuan County	0.129	0.175	0.733
	Emeishan City	0.261	0.262	0.999
	Ebian Yi Autonomous County	0.1	0.148	0.674
	Mabian Yi Autonomous County	0.107	0.162	0.664
Ziyang City	Yanjiang District	0.206	0.207	0.998
	Jianyang City	0.236	0.236	0.998
Mianyang City	Fucheng District	0.554	0.555	1
	Youxian District	0.219	0.219	0.998
	Zitong County	0.12	0.141	0.85
	Santai County	0.204	0.204	0.998
	Yanting County	0.195	0.201	0.969
	Anxian County	0.133	0.14	0.954
	Pingwu County	0.145	0.156	0.93
	Jiangyou City	0.261	0.262	0.998
	Beichuan Qiang Autonomous Region	0.106	0.14	0.756
Zigong City	Ziliujing District	0.368	0.369	0.997
	Yantan District	0.168	0.198	0.849
	Fushun County	0.224	0.225	0.996
Panzhihua City	East District	0.396	0.402	0.984
	West District	0.227	0.564	0.403
	RenHe district	0.127	0.138	0.92
	Miyi County	0.14	0.159	0.883
	Yanbian County	0.199	0.23	0.864
Deyang City	Jingyang District	0.312	0.312	0.998
	Guanghan City	0.205	0.205	0.998
	Shifang City	0.19	0.191	0.998
	Mianzhu City	0.169	0.17	0.996
	Zhongjiang County	0.184	0.185	0.998
	Luoyang County	0.124	0.134	0.92

4 Analysis of the Result

Through the analysis of *CCR*, it appears that Wuhou District in Chengdu and Gongxian County in Yibin have achieved a full effectiveness of comprehensive efficiency (CCR = 1.000) among 147 counties, districts, or cities of Sichuan Province, which indicates that they have obtained a fully efficient outcome. However, two districts, Jinjiang and Chenghua in Chengdu, are in the position of a light level of non-DEA effective degree ($0.700 \leq \text{CCR} < 1.000$). For those of a light level of non-effective degree, the development of electronic administration requires focusing more on the control and management of details and proper distribution of the investment, as there's still some room for improvement. Meanwhile, most of counties (including districts and cities) are in a severe level of non-DEA effective degree ($\text{CCR} \leq 0.500$), which suggests that most counties' e-government development in Sichuan province don't work out. The overall development gap is huge, and most of objects have not achieved the comprehensive effectiveness. The input of the e-government construction has not acquired the desired results. On the one hand, the technical inefficiency leads to the comprehensive ineffectiveness. On the other hand, the low scale efficiency also affects the improvement of the comprehensive efficiency. Therefore, the government should make a huge effort to find out the route to suit the situation, so as to surmount the problem fundamentally. The counties and districts with the lowest CCR, Jinyang County and Butuo County in Liangshan, Rangtang County in Tibetan Qiang Autonomous Prefecture of Aba, Qingshen county in Meishan, Dawu County in Tibetan Autonomous Prefecture of Garze, Aba County in Aba Tibetan and Qiang Autonomous Prefecture, correspond to their relatively weak national economy and social development, as well as the influence of little capital invested by the government on information construction. Thus, problems should be solved and development should be sped up in the future construction of e-government.

The pure technical efficiency results obtained by the BCC model show that Jinjiang district, Wuhou District and Pixian County of Chengdu are with pure technical effective (BCC = 1.00). It indicates that their services systems of e-government run well, and the resources invested can develop productivity efficiently. The pure technical efficiency of 132 counties, districts or cities are lower than 0.5, and it appears that these local governments' technical outcomes are not enough. The last 5 counties, districts or cities with the lowest BCC successively are Hongya County in Meishan, Peng'an County in Nanchong, Qingshen County in Meishan, Pengxi County in Suining, Chaotian District in Guangyuan. It indicates these governments didn't distribute the resources legitimately, and they need more technological innovation to improve the service abilities of e-government from technology level.

From the point of view of scale efficiency, there are 9 counties or districts of Sichuan Province in full scale efficiency. They are Wuhou District, Qingbaijiang District and Shuangliu County in Chengdu, Gongxian County in Yibin, Dachuan District and Xuanhan District in Dazhou, Fucheng District in Mianyang, Lizhou District in Guangyuan and Bazhou District in Bazhong.

These governments input resources in e-government are proper, and returns of scale remain unchanged. The rest 138 counties, districts or cities are not in full scale efficiency, it manifests that these governments waste some resources during the construction of e-government. In the future, the governments need to increase the investment in information technology and science and education, for speeding up the process of urbanization and informatization. What's more, the governments should allocate resources more reasonably to improve the level of government information disclosure, services and websites management. Besides, governments should make more and more people participate in the construction of e-government in order to take better use of the resources.

5 Conclusions

Due to the wave of the informatization and popularization of the network, the construction of e-government in the human condition is getting better and better. Governments at all levels are also vigorously investing money and human resources for the construction of e-government to promote the informatization. To guide the construction of e-government, how to accurately evaluate and scientifically observe the efficiency of e-government is a hotspot and difficulty in the field of current e-government. Although there are a lot of foreign standards being references, different countries may have different national conditions. The real effective e-government construction has to be combined with the actual situation. According to the views of scholars and organizations at home and abroad, based on observation and thinking, we use different indexes and principal component analysis and DEA model to evaluate the e-government construction efficiency of all districts and counties of Sichuan Province. We hope to be able to scientifically evaluate the comprehensive efficiency of various districts and counties of Sichuan Province in the construction of e-government. On the whole, the local government's information department should focus on improving their investment scale and technical efficiency to enhance the efficiency of e-government.

However, this paper also has some limitations. Firstly, we take both input and output into consideration which also have been rotated, extracted and standardized. Therefore, this paper couldn't use the projection algorithm to get the improvement solutions which give the target value and scope for the non-DEA effective. Secondly, DEA focus on the relative efficiency. It means that different evaluation range may have different result. So how to choose the evaluation range is a project worth exploring. In later study, we will commence on diminishing the limitations above and optimize the solution.

Acknowledgements. We would like to express our sincere appreciation to the reviewers for their insightful comments, which have greatly aided us improving the quality of the paper. This research was supported by the Soft Science Research Project of Chengdu (Grant 2016-RK00-00041-ZF).

References

1. Accenture (2006) E-government leadership: high performance and maximum value. http://www.accenture.com
2. CCCT (2015) General report on performance evaluation of Chinese government website in 2015. http://ccct.mofcom.gov.cn/article/zt_jxpg2015
3. Commission E (2006) i2010–a European information society for growth and employment. http://www.docin.com/p-28146066.html
4. CSTC (2013) 2013 Chinese government website performance evaluation index system. http://www.cstc.org.cn/templet/default/show_nbkw.jsp?article_id=123668&id=1619
5. CSTC (2015) 2015 Chinese government website performance evaluation results. http://www.ccidnet.com/2015/1209/10063631.shtml
6. CSTC SPegnoc (2014) The total report of the Sichuan government website performance evaluation 2014
7. Guo X, Zou J (2015) A summary of the current situation and pattern of e-government development in China's rural areas. E-Government 1:90–96 (in Chinese)
8. Kao C (2015) Evaluation and improvement of e-government: the case of European countries. In: International conference on E-democracy & E-government, pp 1211–1216
9. Luna DE et al (2010) Using data envelopment analysis (DEA) to assess government web portals performance. In: Proceedings of the 13th annual international conference on digital government research, pp 107–115
10. Luo D, Shi Y (2010) Evaluation on e-government websites based on rough set and genetic neural network algorithm. In: Proceedings of the third international symposium on computer science and computational technology, pp 124–128
11. Nations U (2014) United nations e-government survey 2014: E-government for the future we want
12. Research C (2004) Description of government website evaluation index system. http://dc.ccw.com.cn
13. Rorissa A, Demissie D, Pardo T (2011) Benchmarking e-government: a comparison of frameworks for computing e-government index and ranking. Gov Inf Q 28:354–362 (in English)
14. Song M, Guan Y (2014) The electronic government performance of environmental protection administrations in Anhui province. Technol Forecast Soc Change 96:79–88 (in Chinese)
15. Song M, Guan Y (2015) Measuring e-government performance of provincial government website in China with slacks-based efficiency measurement. Technol Forecast Soc Change 39:25–31 (in Chinese)
16. Sichuan Provincial Bureau of Statistics SPB (2014) Sichuan province statistical yearbook 2014
17. Zdjelar R (2013) Measuring success of implementation county development strategy by using balanced scorecard method. In: International convention on information & communication technology electronics & microelectronics, pp 1211–1216

Comparing Visitors' Behavior Through Mobile Phone Users' Location Data

Masahide Yamamoto(✉)

Faculty of Foreign Studies, Nagoya Gakuin University, Nagoya, Japan
myama@ngu.ac.jp

Abstract. In recent years, so-called "big data" have been attracting the attention of companies and researchers. This study aims to identify the number of visitors of each period and their characteristics based on the location data of mobile phone users collected by the mobile phone company. The study sites of this survey are tourist destinations in Ishikawa Prefecture and Toyama city, including Kanazawa city, which became nationally popular after the Hokuriku Shinkansen opened in 2015. The opening of the Hokuriku Shinkansen brought more visitors to many areas. However, it also led to fewer visitors in some areas. The positive effect was remarkable in Kanazawa.

Keywords: Mobile phone · Location data · Tourism

1 Introduction

In Japan, many people have struggled to promote tourism in their regions to vitalize their local economies. The low-cost carriers have boosted competition in the transportation industry, and domestic transportation costs have declined in several regions. Therefore, tourism is likely to become increasingly important to local economies. Initially, Japan's tourism industry suffered significant volatility in demand depending on the season and day of the week. In addition, there was significant loss of business opportunities because of congestion during the busy season.

To cope with such volatility, various events have been held to eliminate the seasonal gap. Many events are newly launched. To date, it has been difficult to accurately grasp the extent to which these events attract visitors and the types of people who visit. However, by employing the recently provided Information and Communication Technology (ICT) services, it is possible to verify the number and characteristics of visitors to a particular event.

This study attempts to identify the number of visitors in different periods and their characteristics based on the location data of mobile phone users collected by the mobile phone company. In addition, it also attempts to demonstrate an alternative method to more accurately measure the number of visitors attracted by an event.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_33

2 Method

This study used "MOBILE KUUKAN TOUKEI^{TM}" (mobile spatial statistics) provided by NTT DoCoMo, Inc. and DoCoMo Insight Marketing, Inc. to collect the location data of mobile phone users in order to count the number of visitors at specific tourist destinations and examine their characteristics. MOBILE KUUKAN TOUKEI^{TM} is statistical population data created by a mobile phone network. It is possible to estimate the population structure of a region by gender, age, and residence using this service of a particular company.

Table 1. Survey areas and regional mesh codes

	Survey areas	Regional mesh code	Type of codes
Kanazawa	Kanazawa Station	5436-6591-2	1/2
	Kenrokuen	5436-6572 + 5436-6573-1, 5436-6573-3	Tertiary, 1/2
	Higashi Chayagai	5436-6583-3	1/2
Nanao	Wakura Hot Springs	5536-5703	Tertiary
	Nanao Station	5536-4757	Tertiary
Kaga	Yamanaka Hot Springs	5436-2299, 5436-2390	Tertiary
Wajima	Wajima	5636-0772	Tertiary
Toyama	Toyama Station	5537-0147-1	1/2

Note: A regional mesh code is a code for identifying the regional mesh, which is substantially divided into the same size of a square (mesh) based on the latitude and longitude in order to use it for statistics. The length of one side of a primary mesh is about 80 km, and those of secondary and tertiary meshes are about 10 Km and 1 km respectively.

The sites studied in this survey are tourist destinations in Ishikawa Prefecture and Toyama city, including Kanazawa city, which became nationally popular when the Hokuriku Shinkansen (high-speed railway) opened in 2015. Moreover, the locations and characteristics of the individuals obtained herein are derived through a non-identification process, aggregation processing, and concealment processing. Therefore, it is impossible to identify specific individuals.

The survey areas are presented in Table 1 and Fig. 1. A regional mesh code is a code for identifying the regional mesh. It stands for an encoded area that is substantially divided into the same size of a square (mesh) based on the latitude and longitude in order to use it for statistics. With regard to regional mesh, there are three types of meshes: primary, secondary, and tertiary. The length of one side of a primary mesh is about 80 km, and those of secondary and tertiary meshes are about 10 Km and 1 km respectively.

In addition, split regional meshes also exist, which are a more detailed regional division. A half-regional mesh is a tertiary mesh that is divided into two equal pieces in the vertical and horizontal directions. The length of one side is about 500 m. Furthermore, the length of one side of a quarter and 1/8 regional meshes is about 250 m and 125 m respectively.

Fig. 1. Survey areas

This study analyzed the location data collected from NTT DoCoMo, Inc. to consider the effect of the opening of the Hokuriku Shinkansen on the survey areas.

3 Previous Research

Previous tourism marketing research has primarily focused on the ways service promises are made and kept, mostly generating frameworks to improve managerial decisions or providing insights on associations between constructs [4]. Big data have become important in many research areas, such as data mining, machine learning, computational intelligence, information fusion, the semantic Web, and social networks [3]. To date, several attempts have been made to use large-scale data or mobile phone location data in tourism marketing studies.

Most studies dealing with big data in tourism were published after 2010. Fuchs et al. [5] presented a knowledge infrastructure that has recently been implemented at the leading Swedish mountain tourism destination, Åre. Using a Business Intelligence approach, the Destination Management Information System Åre (DMIS-Åre) drives knowledge creation and application as a precondition of organizational learning at tourism destinations. Xiang et al. [9] tried to apply big data to tourism marketing. The study aimed to explore and demonstrate the utility of big data analytics to better understand important hospitality issues, namely, the relationship between hotel guest experience and satisfaction. Specifically, the investigators applied a text analytical approach to a large number of consumer reviews extracted from Expedia.com to deconstruct hotel guest experiences and examine the association with satisfaction ratings. These studies are similar to this study in that they attempted to utilize big data in tourism. However, the research methods and objectives of these studies are different from that of the present study.

Studies on using mobile phone location data for tourism surveys can be traced back to 2008. Ahas et al. [2] introduced the applicability of passive mobile positioning data for studying tourism. They used a database of roaming location (foreign phones) and call activities in network cells: the location, time, random identification, and country of origin of each called phone. Using examples from Estonia, their study described the peculiarities of the data, data gathering, sampling, the handling of the spatial database, and some analytical methods to demonstrate that mobile positioning data have valuable applications for geographic studies. Japan Tourism Agency conducted a similar study using international roaming service in December 2014 [1].

Since the creative work of Ahas et al. [2], several studies employing location data have emerged. Liu et al. [7] investigated the extent to which behavioral routines could reveal the activities being performed at mobile phone call locations captured when users initiate or receive voice calls or messages. Using data collected from the natural mobile phone communication patterns of 80 users over more than a year, they assessed the approach via a set of extensive experiments. Based on the ensemble of models, they achieved prediction accuracy of 69.7%. The experiment results demonstrated the potential to annotate mobile phone locations based on the integration of data mining techniques with the characteristics of underlying activity-travel behavior.

Alternative related studies have also been conducted. Gao and Liu [6] attempted to examine the methods used to estimate traffic measures using information from mobile phones, accounting for the fact that each vehicle likely contains more than one phone because of the popularity of mobile phones. Steenbruggen et al. [8] used mobile phone data to provide new spatio-temporal tools for improving urban planning and reducing inefficiencies in current urban systems. They addressed the applicability of such digital data to develop innovative applications to improve urban management.

As described above, this study surveyed previous related research. Among those studies, the present study could be characterized as similar to Ahas et al. [2]. However, Ahas et al. [2] is based on results obtained by analyzing data roaming activity. Mobile phone users in the study are obviously limited. Therefore, whether the knowledge gained applies to the average traveler is not clear. In the present study, I analyzed data provided by NTT DoCoMo, Inc., which is the largest mobile phone service provider in Japan. Therefore, their data should be more reliable in that the parameter is quite large.

4 Results

In general, the number of visitors has been increasing since the Hokuriku Shinkansen was launched on May 14, 2015, with the exception of Nanao station.

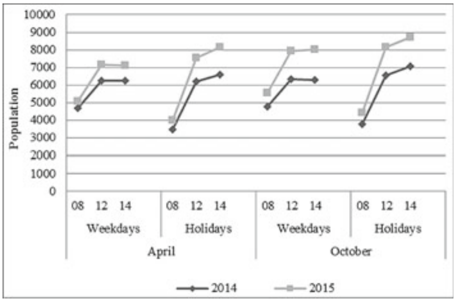


Fig. 2. Visitor transitions at Kanazawa station

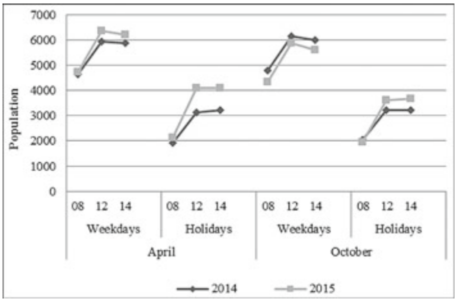


Fig. 3. Visitor transitions at Toyama station

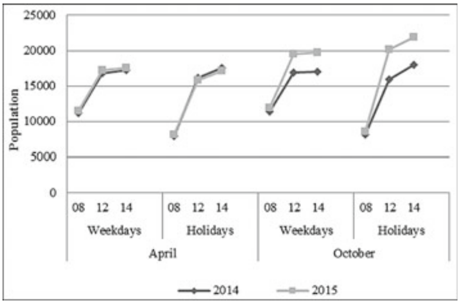


Fig. 4. Visitor transitions at Kenrokuen

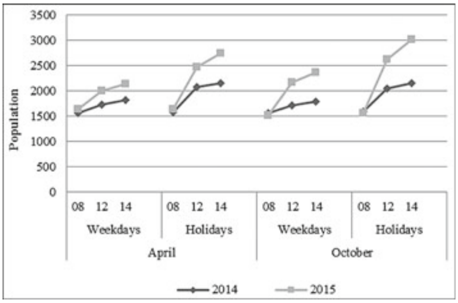


Fig. 5. Visitor transitions at Higashi Chayagai

It should be noted that these "visitor" also include the residents living there, because the data cannot exclude them. Of course, I tried to exclude residential areas as much as possible when I specified the regional mesh codes. However, it was rather difficult to do that, because the mesh codes are square-shaped.

First, I compared the results of two larger cities, Kanazawa and Toyama (see Figs. 2 and 3). Both these cities have a station at which the Hokuriku Shinkansen stops. It should be noted that Kanazawa city and Toyama attracted more visitors in the afternoons, whereas Wakura and Wajima, which are located on the Noto Peninsula, had more visitors in the mornings (8:00 a.m.–9:00 am). Visitors to Toyama Station demonstrated approximately the same trend as those visiting Kanazawa Station. However, there were fewer visitors on holidays than on weekdays in Toyama.

I then examined the data of Kanazawa city (see Figs. 4 and 5). There were three survey areas in this city: Kanazawa station, Kenrokuen Park and Higashi Chayagai.

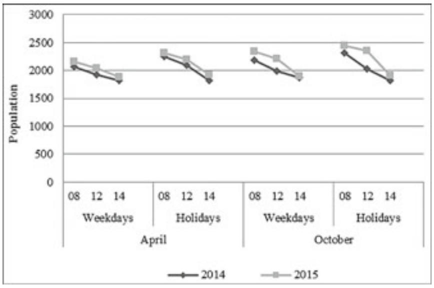


Fig. 6. Visitor transitions at Wajima

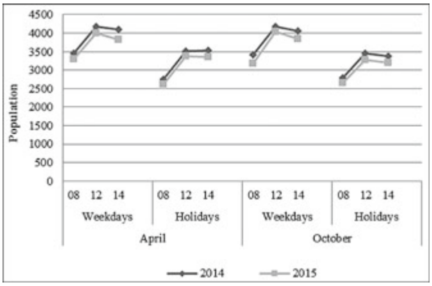


Fig. 7. Visitor transitions at Nanao station

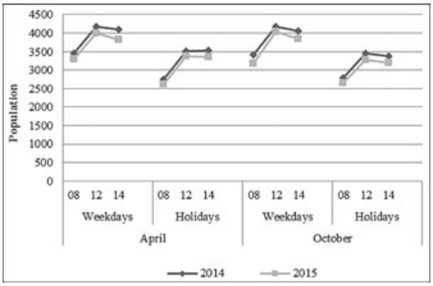


Fig. 8. Visitor transitions at Wakura hot springs

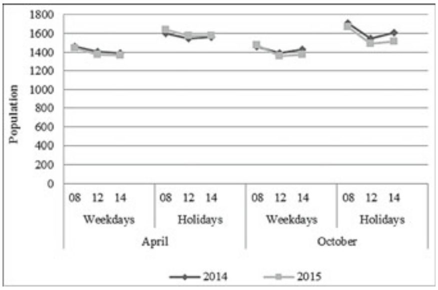


Fig. 9. Visitor transitions at Yamanaka hot springs

Wajima is famous for its Asaichi (morning market). Presumably, visitors enjoy shopping at the Noto Shokusai market near Nanao Station during the daytime, move on to Wakura hot springs later in the day, and then return to the morning market the next day (see Figs. 6, 7 and 8).

Wajima is also the setting for a TV drama "Mare", which was broadcasted nationwide from April to September 2015. Visits to Wajima have slightly increased in 2015. The increase in visits could be attributed to this TV drama rather than the opening of the Hokuriku Shinkansen, because Nanao attracted fewer visitors in 2015 than in 2014 in spite of being a better location and nearer to Kanazawa.

Regarding Yamanaka hot springs, there was no significant difference in visits between different periods. Some visitors might have spent more than one night in Yamanaka hot springs (see Fig. 9).

Although both Wakura and Yamanaka hot springs are a little far from Kanazawa, their results were contrary. The TV drama might have increased the number of tourists in Wakura (Figs. 10 and 11).

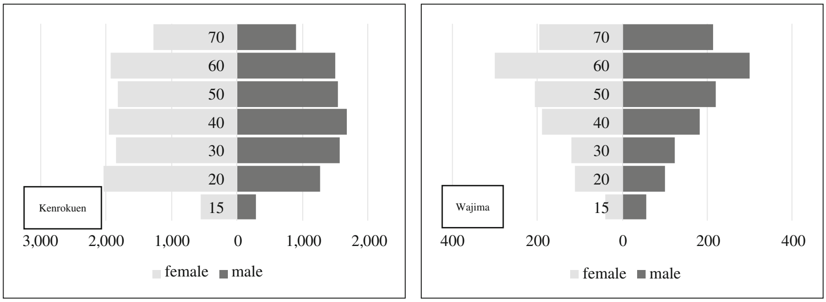


Fig. 10. Visitors' gender distribution at Kenrokuen and Wajima (12:00 a.m.–1:00 p.m. on holidays in October 2015)

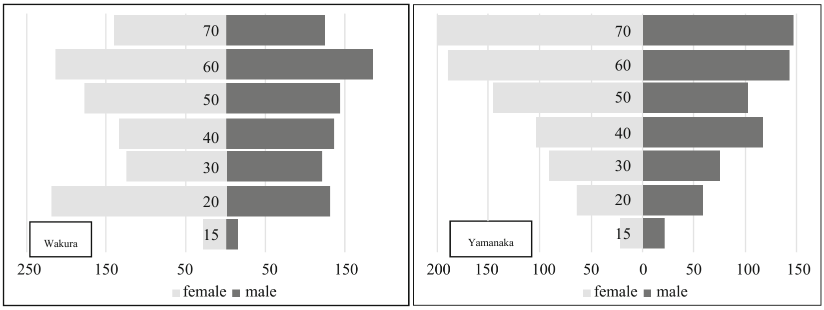


Fig. 11. Visitors' gender distribution at Wakura hot springs and Yamanaka hot springs (12:00 a.m.–1:00 p.m. on holidays in October 2015)

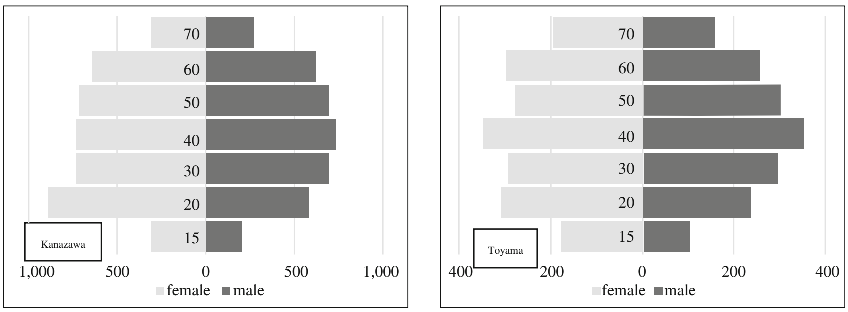


Fig. 12. Visitors' gender distribution at Kanazawa station and Toyama Station (12:00 a.m.–1:00 p.m. on holidays in October 2015)

Kanazawa city (particularly Kenrokuen) attracted a variety of visitors, including many female visitors.

On the other hand, many elderly people (over 60 years old) visited the hot springs and the Noto Peninsula. I presume that the local people account for a large proportion of these visitors.

A comparison of visitors at Kanazawa Station with those at Toyama Station found that Kanazawa Station attracted visitors from a wider area of Japan. Mobile phone users around Kanazawa Station were from 235 municipalities, including Ishikawa Prefecture, whereas those at Toyama Station were from 43 municipalities, including Toyama Prefecture (12:00 a.m.–1:00 p.m. on holidays in October 2015) (Fig. 12).

Although the Hokuriku Shinkansen stops at both stations, the results suggest that Kanazawa has been more successful so far in attracting visitors (see Table 2). The number of visitors from Tokyo (gray column) increased in both the cities.

Table 2. Visitors' residential distribution at Kanazawa Station and Toyama Station(12:00 a.m.–1:00 p.m. on holidays in October 2014 and 2015)

Kanazawa station			Toyama station		
Residence	2015	2014	Residence	2015	2014
Toyama, Toyama Pref.	131	147	Kanazawa, Ishikawa Pref.	67	60
Takaoka, Toyama Pref.	93	94	Fukui, Fukui Pref.	18	13
Fukui, Fukui Pref.	91	80	Setagaya-ku, Tokyo	16	n/a
Myoko, Niigata Pref.	47	n/a	Suginami-ku, Tokyo	15	n/a
Nanto, Toyama Pref.	34	31	Yamagata, Yamagata Pref.	14	n/a
Setagaya-ku, Tokyo	34	17	Takayama, Gifu Pref.	14	n/a
Imizu, Toyama Pref.	34	42	Nakamura-ku, Nagoya, Aichi Pref.	13	n/a
Oyabe, Toyama Pref.	33	39	Minami-ku, Niigata, Niigata Pref.	13	n/a
Sakai, Fukui Pref.	32	30	Gifu, Gifu Pref.	12	n/a
Omachi, Nagano Pref.	30	n/a	Hakusan, Ishikawa Pref.	12	20
Ota-ku, Tokyo	28	12	Ota-ku, Tokyo	12	n/a
Tsubame, Niigata Pref.	28	n/a	Himeji, Hyogo Pref.	12	n/a
Bunkyo-ku, Tokyo	27	n/a	Nagano, Nagano Pref.	12	n/a
Nagano, Nagano Pref.	24	n/a	Nerima-ku, Tokyo	11	n/a
Hiratsuka, Kanagawa Pref.	24	n/a	Shinagawa-ku, Tokyo	11	n/a
Sanda, Hyogo Pref.	23	n/a	Hida City, Gifu Pref.	11	10
Tonami, Toyama Pref.	23	26	Joetsu, Niigata Pref.	11	n/a
Suginami-ku, Tokyo	22	16	Kawaguchi, Saitama Pref.	10	n/a
Nerima-ku, Tokyo	22	10	Adachi-ku, Tokyo	10	n/a
Tsu, Mie Pref.	22	n/a	Takatsuki, Osaka	10	n/a

Note: The figures above do not include visitors from Ishikawa Pref. for Kanazawa Station and Toyama Pref. for Toyama Station. Numbers less than ten are represented as "n/a"

Despite the fact that Toyama is nearer to Tokyo than Kanazawa, the latter successfully attracted more visitors from Tokyo.

Regarding the two prefectures, Ishikawa and Toyama, they both do have many wonderful tourist attractions. However, as for the two cities, it seems that Kanazawa is more attractive to tourists.

5 Conclusion

This study attempted to identify the number of visitors at two points in time at various places in Japan and their characteristics using the location data of mobile phone users collected by the mobile phone company.

As explained above, the opening of the Hokuriku Shinkansen increased the number of visitors to many areas. However, it also led to fewer visitors in some other areas. Its positive effect was remarkable in Kanazawa.

Numerous events have been recently held in Japan to attract visitors. In addition to using the "MOBILE KUUKAN TOUKEI^{TM}", combining it with other ICT services, such as Google Trends, can help better predict the number of visitors at new events. Specifically, by combining the "MOBILE KUUKAN TOUKEI^{TM}" and the transition of the search results for a particular tourist destination, it would be possible to more accurately predict the number of tourists. If we can realize more accurate demand forecasting, it would be possible to optimize the necessary goods and number of non-regular employees in advance. Moreover, understanding consumers' characteristics beforehand could enable us to optimize the services, which could influence customer satisfaction.

Acknowledgements. This work was supported by JSPS KAKENHI Grant Number JP15K01970. The "MOBILE KUUKAN TOUKEI^{TM}" and logo are trademarks of NTT DOCOMO, Inc.

References

1. Agency JT (2014) Keitaidenwa kara erareru ichijouhou tou wo katuyousita hounichi gaikokujin doutaichousa houkokusho [foreign visitors dynamics research report utilizing mobile phone location information]. Technical report, Japan Tourism Agency
2. Ahas R, Aasa A et al (2008) Evaluating passive mobile positioning data for tourism surveys: an estonian case study. Tourism Manage 29(3):469–486
3. Bello-Orgaz G, Jung JJ, Camacho D (2016) Social big data: recent achievements and new challenges. Inf Fusion 28:45–59
4. Dolnicar S, Ring A (2014) Tourism marketing research: past, present and future. Ann Tourism Res 47:31–47
5. Fuchs M, Höpken W, Lexhagen M (2014) Big data analytics for knowledge generation in tourism destinations - a case from Sweden. J Destination Mark Manage 3(4):198–209
6. Gao H, Liu F (2013) Estimating freeway traffic measures from mobile phone location data. Eur J Oper Res 229(1):252–260
7. Liu F, Janssens D et al (2013) Annotating mobile phone location data with activity purposes using machine learning algorithms. Expert Syst Appl 40(8):3299–3311

8. Steenbruggen J, Tranos E, Nijkamp P (2015) Data from mobile phone operators: a tool for smarter cities? Telecommun Policy 39(3–4):335–346
9. Xiang Z (2015) What can big data and text analytics tell us about hotel guest experience and satisfaction? Int J Hosp Manage 44(44):120–130

Research on Geo/Geo/1 Retrial Queue with Working Vacation Interruption and Nonpersistent Customers

Mingcong Wu[1], Yong Huang[1](✉), Yang Song[2], Liang Zhao[1], and Jian Liu[1]

[1] Business School, Sichuan University, Chengdu 610065, People's Republic of China
huangyong_scu@aliyun.com

[2] Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, People's Republic of China

Abstract. Queueing phenomenon is prevalent in all areas of society and a reasonable queueing design can improve service efficiency and customer satisfaction. In order to adapt to the new requirements of web service system, this paper studies a Geo/Geo/1 retrial queue with working vacation interruption and nonpersistent customers. Firstly, a series of assumptions about the queueing system are put forward and the corresponding transition probability matrix is obtained. Then the stationary condition of the queueing system is derived. After that, the stationary distribution and the performance measures are obtained by using the matrix-analytic method. Finally, numerical analysis is conducted to discuss the effect of parameters on performance measures, furthermore, the performance of the queueing model is optimized to obtain the best parameters and the minimum cost.

Keywords: Web service · Discrete-time queueing system · Retrial · Working vacation interruption · Nonpersistent customers

1 Introduction

With the development of the internet, the internet-based web service industry is growing explosively, and web service of higher quality is required. The basic application structure of web service is shown in Fig. 1. In the system, users can request services according to their own needs through the internet. After their service requests are passed to web server, the web server will interact with application server and database server for different requests, and process various requests properly. Finally, the results of the requests will be returned to the users. The function of the application server is to guarantee the business logic in the activities and coordinates the information exchange among the users. The task of the database server is to complete query, storage, modify and other functions of web service data in the database.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_34

Web service system needs to use the adapted queueing model in dealing with the order of the users, but the traditional queueing model cannot accurately deal with this scene. When the arriving user finds the server is temporarily unable to provide service, he will leave the server and repeat requests after a while, which is called retrial strategy. Considering the utilization of server, managers often allow server to serve at a relatively low rate when there are no users in the system, but when the number of the users requesting service increases in the system, the service rate needs to return to a higher level immediately, which is called working vacation interruption strategy. Besides, users who conduct business on the internet tend to give up receiving service and leave the service system because they don't want to wait too long, which is called nonpersistent customers strategy. Scholars have studied the above three cases separately, but there are no studies of the queueing model which possess these three strategies at the same time. This paper tries to study the combination of these three strategies, and solve the construction of the queueing model of web service system under these conditions.

This paper is organized as follows. In Sect. 1, relevant literatures are reviewed. Section 2 is dedicated to formulate a queueing model to solve the problem mentioned above. In Sect. 3, a series of numerical analyses are presented and the performance of the model is optimized. Section 4 gives some conclusions.

2 Literature Review

Queueing theory studies the working process of random service system. It originated from the telephone conversation in the early 20^{th} century. In 1909–1920, Erlang, the Danish mathematician and electrical engineer, studied the problem of telephone conversation using the method of probability theory, which created this subject of applied mathematics. Queueing system is divided into continuous-time queueing system and discrete-time queueing system. Compared with the former, the research of discrete-time queueing system started relatively late and

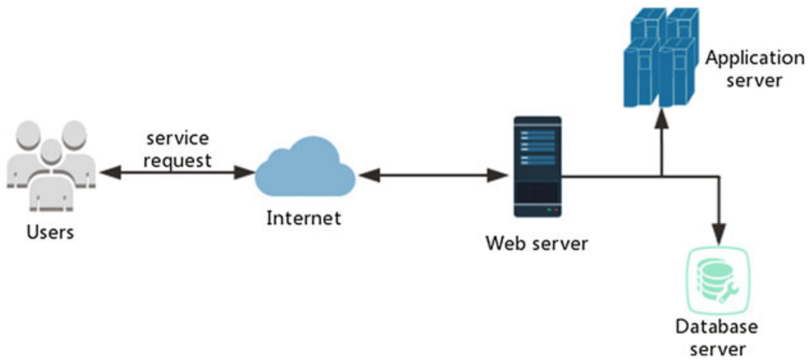


Fig. 1. The basic application structure of web service

the early research results about the discrete-time queueing system are few. But with the development of computer and communication in which time is slotted, the discrete-time queueing system has become one of the research hotspots of queueing theory in recent years. Meisling [10] made a pioneering work on the discrete-time queueing system, and then a series of studies have deepened and developed the analysis of this classical model. Kobayash and Konheim [3] studied the application of the discrete-time queueing system in the field of computer communication network, and pointed out that the discrete-time queueing system is more suitable for the modeling and analysis of computer networks, which has played a positive role in the research and application of the discrete-time queueing system.

Retrial queueing system is an important part of queueing theory research. The characteristic of retrial queueing system is that customer who requests service will enter a queue called orbit and retry after a random time when he finds all servers are busy. Kosten [4] first proposed a retry service system. Yang and Li [18] were the first to study a discrete-time retrial queue, they studied the steady-state queue size distribution of the discrete-time Geo/G/1 retrial queue. Then retrial queueing system is widely used in telephone switching systems, computer networks and communications networks. So far, these research areas are still vibrant. Some of the latest developments regarding retrial queueing systems are Dimitriou [2], Rajadurai [13] and so on. If the server is occupied, the customer retrying from the orbit may leave the queueing system or return to the orbit, which is called nonpersistent customers phenomenon. Palm [11] first proposed the issue of nonpersistent customers. Then, many scholars introduced nonpersistent customers into various queueing systems. Liu and Song [9] introduced nonpersistent customers to a Geo/Geo/1 retrial queue with working vacations. Phung-Duc [12] introduced two types of nonpersistent customers into a M/M/c/K ($K \geq c \geq 1$) retrial queues. These models are also widely used in real life. In fact, the introduction of nonpersistent customers is of great significance to imitate the impatience characteristic of customers in reality.

The queueing systems with vacation and working vacation have always been a research hotspot. Queueing system with vacation is an extension of the classical queueing system, which allows server not to serve customers at certain time called vacation. For more information about queueing system with vacation, see Tian and Zhe [16]. If server does not completely stop serving on vacation, but rather serve at a slower service rate than the rate of regular busy period, this vacation policy is called working vacation. Servi and Finn [14] first introduced the queueing system with working vacation when they studied the model of communication network. They introduced the working vacation mechanism into the M/M/1 queue. Then, many scholars introduced working vacation into various queueing system. Tian et al. [17] studied a discrete-time queueing system with working vacation. Li and Tian [7] studied a discrete-time Geo/Geo/1 queue with single working vacation. On the basis of working vacations, working vacation interruption strategy was introduced, which means that once a certain indicator (such as the number of customers in the system) reaches a certain value on working vacation, server can stop working vacation and return to normal service level immediately. Li and Tian [6] first proposed working vacation interruption strategy. They introduced working vacation and vacation interruption into M/M/1 queue, then they studied the corresponding discrete-time queue [5]. Recently, studies on working vacation interruption are continued. Li et al. [8] introduced working vacation interruption into discrete-time Geo/Geo/1 retrial queue, Gao et al. [15] studied an M/G/1 queue with single working vacation and vacation interruption under Bernoulli schedule.

Previous studies can't address the real-world scenes of web service that are mentioned. So, this paper considers introducing working vacation interruption and nonpersistent customers to discrete-time Geo/Geo/1 retrial queue to provide the solution.

3 Method

In order to deal with retrial, working vacation interruption and nonpersistent customers phenomena in web service, this section constructs an adapted queueing model and studies the queueing system. Section 3.1 put forwards a series of assumptions about the queueing model (for convenience of study, assume that the number of server is one) and obtains the corresponding transition probability matrix for subsequent derivation. The research of queueing system needs to be based on the stability of the system, so Sect. 3.2 derives the condition under which the system is stable, and the rate matrix R for the derivation of stationary distribution. The main purpose of the research of queueing system is to study the states and the indexes when the system is stable, so Sect. 3.3 derives the stationary distribution and some performance measures using matrix-analytic method.

3.1 Model

The Geo/Geo/1 retrial queue with working vacation interruption and nonpersistent customers is assumed as follows:

First of all, the rules of service and customer arrival are assumed. Assume the beginning and ending of service occur at slot division point $t = n, n = 0, 1, \cdots$. The service time S_b in a regular busy period follows a geometric distribution with parameter μ_b. The service time S_v in a working vacation period follows a geometric distribution with parameter μ_v. Note that $0 < \mu_v < \mu_b < 1$. Assume customer arrivals occur at the slot $t = n, n = 0, 1, \cdots$. Inter-arrival time, which is an independent and identically distributed sequence, follows a geometric distribution with parameter during a regular busy period and a geometric distribution with parameter λ_b during a working vacation period. Also Note that $0 < \lambda_v < \lambda_b < 1$.

Then, assume the rules of retrial and nonperisistent customers strategies. Customer from the orbit of infinite size requests retrial at the slot $t = n, n = 0, 1, \cdots$. And the time between two successive retrials follows a geometric distribution with parameter α. If the customer arrival and retrial occur at the same instant when the server is not occupied, we assume that the arriving customer receives the service. And, due to the introduction of nonpersistent customers strategy, suppose the probability that the retrial customer leaves system is $q(0 < q < 1)$, and the probability that retrial customer returns to the orbit is $1 - q$.

Finally, working vacation and vacation interruption are assumed. The server begins a working vacation each time when there is no customer in the system, i.e., the server is free and there is no customer in the orbit. Assume the beginning and ending of working vacation occur at the slot $t = n^+, n = 0, 1, \cdots$. At the end of each working vacation, the server begins a new vacation only if there is no customer in the system. And the vacation time V follows a geometric distribution with parameter θ, $(0 < \theta < 1)$. If the service is complete at slot t = n$^-$ and there is customer in the system at slot $t = n^+$, the server will stop the working vacation immediately and return to the normal service level, i.e., working vacation interruption strategy.

In this paper, for any real number $x \in [0, 1]$, denote $\overline{x} = 1 - x$. Assume that interarrival time, service time, and working vacation time are mutually independent.

After completing the above assumptions, the occurring order of the random events can be described in Fig. 2 using number axis.

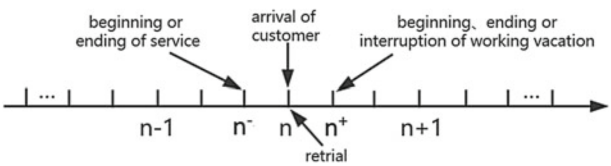


Fig. 2. The occurring order of the random events

Meanwhile, the structure of the queueing model in web service can be shown in Fig. 3. User accesses the web server through the internet. He joins the orbit if the web server is busy, or else he receives service immediately and leaves the system after the service is completed. The retrial user from the orbit receives service immediately if the web server is free, otherwise he may leave the system or come back to the orbit.

Let Q_n be the number of customers in the orbit at the slot n^+, and J_n the state of server at the slot n^+. There are four possible states of the server as follows:

(1) The state $J_n = 0$ denotes the server is free in a working vacation period at the slot n^+.
(2) The state $J_n = 1$ denotes the server is busy in a working vacation period at the slot n^+.
(3) The state $J_n = 2$ denotes the server is free in a regular busy period at the slot n^+.

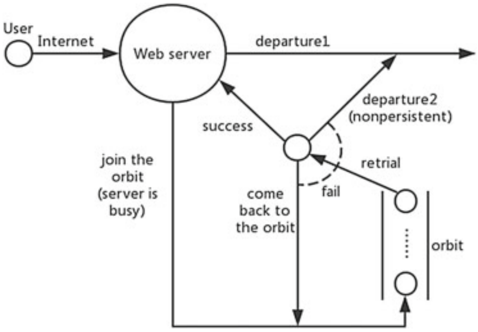


Fig. 3. The structure of the queueing model in web service

(4) The state $J_n = 3$ denotes the server is busy in a regular busy period at the slot n^+.

Then, $\{Q_n, J_n\}, n \geq 0$, is a Markov chain with state space $\Omega = \{(k, j), k \geq 0, j = 0, 1, 2, 3\}$. Note that, the probability that the server is free in a regular busy period is zero when there is no customer in the orbit.

Using the lexicographical sequence for the states, the transition probability matrix of $\{Q_n, J_n\}$ can be written as:

$$P = \begin{pmatrix} B_1 & B_0 & & \\ A_2 & A_1 & A_0 & \\ & A_2 & A_1 & A_0 \\ & & \ddots & \ddots & \ddots \end{pmatrix},$$

where

$$B_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \bar{\mu}_v \lambda_v \bar{\theta} & 0 & \bar{\mu}_v \lambda_v \theta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{\mu}_b \lambda_b \end{pmatrix}, B_1 = \begin{pmatrix} \bar{\lambda}_v & \lambda_v \bar{\theta} & 0 & \lambda_v \theta \\ \mu_v \bar{\lambda}_v & \bar{\mu}_v \bar{\lambda}_v \bar{\theta} & 0 & \mu_v \lambda_v + \bar{\mu}_v \bar{\lambda}_v \theta \\ 1 & 0 & 0 & 0 \\ \mu_b \bar{\lambda}_b & 0 & 0 & \mu_b \lambda_b + \bar{\mu}_b \bar{\lambda}_b \end{pmatrix},$$

$$A_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \bar{\mu}_v \lambda_v (\bar{\alpha} + \alpha \bar{q}) \bar{\theta} & 0 & \bar{\mu}_v \lambda_v (\bar{\alpha} + \alpha \bar{q}) \theta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{\mu}_b \lambda_b (\bar{\alpha} + \alpha \bar{q}) \end{pmatrix},$$

$$A_1 = \begin{pmatrix} \bar{\lambda}_v \bar{\alpha} \bar{\theta} & \lambda_v \alpha \bar{q} \bar{\theta} + \lambda_v \bar{\alpha} \bar{\theta} & \bar{\lambda}_v \bar{\alpha} \theta & \lambda_v \bar{\alpha} \theta + \lambda_v \alpha \bar{q} \theta \\ 0 & \bar{\mu}_v \bar{\lambda}_v (\bar{\alpha} + \alpha \bar{q}) \bar{\theta} + \bar{\mu}_v \lambda_v \alpha q \bar{\theta} & \mu_v \bar{\lambda}_v \bar{\alpha} & (\mu_v \lambda_v + \bar{\mu}_v \bar{\lambda}_v \theta)(\bar{\alpha} + \alpha \bar{q}) + \bar{\mu}_v \lambda_v \alpha q \theta \\ 0 & 0 & \bar{\lambda}_b \bar{\alpha} & \lambda_b (\bar{\alpha} + \alpha \bar{q}) \\ 0 & 0 & \mu_b \bar{\lambda}_b \bar{\alpha} & (\mu_b \lambda_b + \bar{\mu}_b \bar{\lambda}_b)(\bar{\alpha} + \alpha \bar{q}) + \bar{\mu}_b \lambda_b \alpha q \end{pmatrix},$$

$$A_2 = \begin{pmatrix} 0 & \lambda_v \alpha q \bar{\theta} + \bar{\lambda}_v \alpha \bar{\theta} & 0 & \lambda_v \alpha q \theta + \bar{\lambda}_v \alpha \theta \\ 0 & \bar{\mu}_v \bar{\lambda}_v \alpha q \bar{\theta} & 0 & (\bar{\mu}_v \bar{\lambda}_v \theta + \mu_v \lambda_v) \alpha q + \mu_v \bar{\lambda}_v \alpha \\ 0 & 0 & 0 & \lambda_b \alpha q + \bar{\lambda}_b \alpha \\ 0 & 0 & 0 & (\bar{\mu}_b \bar{\lambda}_b + \mu_b \lambda_b) \alpha q + \mu_b \bar{\lambda}_b \alpha \end{pmatrix}.$$

It can be seen from the block structure of the transition probability matrix that is a quasi-birth and death (QBD) process.

3.2 Stability Condition and Rate Matrix R

In this section, after using the matrices A_0, A_1 and A_2 to get the matrix A, derive the stability condition of the queueing system using the matrix A.
Let $A = A_0 + A_1 + A_2$, then,

$$A = \begin{pmatrix} \bar{\lambda}_v \bar{\alpha} \bar{\theta} & \alpha \bar{\theta} + \lambda_v \bar{\alpha} \bar{\theta} & \bar{\lambda}_v \bar{\alpha} \theta & \alpha \theta + \lambda_v \bar{\alpha} \theta \\ 0 & \bar{\mu}_v \bar{\theta} & \mu_v \bar{\lambda}_v \bar{\alpha} & \bar{\mu}_v \theta + \mu_v \lambda_v + \mu_v \bar{\lambda}_v \alpha \\ 0 & 0 & \bar{\lambda}_b \bar{\alpha} & \lambda_b + \bar{\lambda}_b \alpha \\ 0 & 0 & \mu_b \bar{\lambda}_b \bar{\alpha} & \bar{\mu}_b + \mu_b \lambda_b + \mu_b \bar{\lambda}_b \alpha \end{pmatrix}.$$

Theorem 1. *The QBD process $\{Q_n, J_n\}$ is positive recurrent (stable) if and only if $\bar{\mu}_b (\bar{\lambda}_b \bar{\alpha} - 1)(\alpha q - \lambda_b) < \mu_b (\bar{\lambda}_b + \lambda_b q) \alpha$.*

Proof. Since A is reducible, according to Theorem 1 in Latouche and Ramaswami [1], know that the QBD process is positive recurrent if and only if

$$v \begin{pmatrix} 0 & \lambda_b \alpha q + \bar{\lambda}_b \alpha \\ 0 & \bar{\mu}_b \bar{\lambda}_b \alpha q + \mu_b \bar{\lambda}_b \alpha + \mu_b \lambda_b \alpha q \end{pmatrix} e > v \begin{pmatrix} 0 & 0 \\ 0 & \bar{\mu}_b \lambda_b (\bar{\alpha} + \alpha \bar{q}) \end{pmatrix} e,$$

where v is the stationary probability vector of

$$\begin{pmatrix} \bar{\lambda}_b \bar{\alpha} & \lambda_b + \bar{\lambda}_b \alpha \\ \mu_b \bar{\lambda}_b \bar{\alpha} & \bar{\mu}_b + \mu_b \lambda_b + \mu_b \bar{\lambda}_b \alpha \end{pmatrix},$$

e is a column vector with all elements equal to one. By solving

$$\begin{cases} v \begin{pmatrix} \bar{\lambda}_b \bar{\alpha} & \lambda_b + \bar{\lambda}_b \alpha \\ \mu_b \bar{\lambda}_b \bar{\alpha} & \bar{\mu}_b + \mu_b \lambda_b + \mu_b \bar{\lambda}_b \alpha \end{pmatrix} = v \\ ve = 1 \end{cases},$$

get

$$v = \left(\frac{\mu_b \bar{\lambda}_b \bar{\alpha}}{1 + \mu_b \bar{\lambda}_b \bar{\alpha} - \bar{\lambda}_b \bar{\alpha}}, \frac{1 - \bar{\lambda}_b \bar{\alpha}}{1 + \mu_b \bar{\lambda}_b \bar{\alpha} - \bar{\lambda}_b \bar{\alpha}}\right).$$

Then take v into the inequation above, after a series of algebraic operations, the QBD process $\{Q_n, J_n\}$ is positive recurrent if and only if $\bar{\mu}_b(\bar{\lambda}_b\bar{\alpha}-1)(\alpha q-\lambda_b) < \mu_b(\bar{\lambda}_b+\lambda_b q)\alpha$, i.e., the stability condition of the queueing system.

Next, before deriving the stationary distribution of the system, need to obtain the minimal non-negative solution R which satisfies the matrix quadratic equation

$$R = R^2A_2 + RA_1 + A_0. \tag{1}$$

Theorem 2. *If* $\bar{\mu}_b(\bar{\lambda}_b\bar{\alpha}-1)(\alpha q-\lambda_b) < \mu_b(\bar{\lambda}_b+\lambda_b q)\alpha$*, the matrix equation above has the minimal non-negative solution*

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & r_1 & r_2 & r_3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & r_4 & r_5 \end{pmatrix},$$

where

$$r_1 = \frac{1-\bar{\mu}_v\bar{\lambda}_v(\bar{\alpha}+\alpha\bar{q})\bar{\theta}-\bar{\mu}_v\lambda_v\alpha q\bar{\theta}-\sqrt{D}}{2\bar{\mu}_v\bar{\lambda}_v\alpha q\bar{\theta}}, \; r_2 = \frac{\mu_v\bar{\lambda}_v\bar{\alpha}r_1+\mu_b\bar{\lambda}_b\bar{\alpha}r_3}{1-\bar{\lambda}_b\bar{\alpha}},$$

$$r_3 = \frac{Mr_1{}^2+Nr_1+\bar{\mu}_v\lambda_v(\bar{\alpha}+\alpha\bar{q})\theta(1-\bar{\lambda}_b\bar{\alpha})}{(1-\bar{\lambda}_b\bar{\alpha})[1-L-F(r_1+r_5)-(\lambda_b\alpha q+\bar{\lambda}_b\alpha)r_4]-G},$$

$$r_4 = \frac{\mu_b\bar{\mu}_b\lambda_b\bar{\lambda}_b(\bar{\alpha}+\alpha\bar{q})\bar{\alpha}}{\bar{\mu}_b\bar{\lambda}_b\alpha q-\bar{\mu}_b\bar{\lambda}_b\alpha\bar{\alpha}q+\mu_b\bar{\lambda}_b\alpha+\mu_b\lambda_b\alpha q},$$

$$r_5 = \frac{\bar{\mu}_b\lambda_b(\bar{\alpha}+\alpha\bar{q})(1-\bar{\lambda}_b\bar{\alpha})}{\bar{\mu}_b\bar{\lambda}_b\alpha q-\bar{\mu}_b\bar{\lambda}_b\alpha\bar{\alpha}q+\mu_b\bar{\lambda}_b\alpha+\mu_b\lambda_b\alpha q},$$

$$D = [\bar{\mu}_v\bar{\lambda}_v(\bar{\alpha}+\alpha\bar{q})\bar{\theta}+\bar{\mu}_v\bar{\lambda}_v\alpha q\bar{\theta}-1]^2-4\bar{\mu}_v{}^2\bar{\lambda}_v\lambda_v\alpha q(\bar{\alpha}+\alpha\bar{q})\bar{\theta}^2,$$

$$M = (\bar{\mu}_v\bar{\lambda}_v\alpha q\theta+\mu_v\bar{\lambda}_v\alpha+\mu_v\lambda_v\alpha q)(1-\bar{\lambda}_b\bar{\alpha})+\mu_v\bar{\lambda}_v\bar{\alpha}(\lambda_b\alpha q+\bar{\lambda}_b\alpha),$$

$$N = [(\mu_v\lambda_v+\bar{\mu}_v\bar{\lambda}_v\theta)(\bar{\alpha}+\alpha\bar{q})+\bar{\mu}_v\lambda_v\alpha q\theta](1-\bar{\lambda}_b\bar{\alpha})+\mu_v\bar{\lambda}_v\lambda_b(\bar{\alpha}+\alpha\bar{q})\bar{\alpha},$$

$$L = (\mu_b\lambda_b+\bar{\mu}_b\bar{\lambda}_b)(\bar{\alpha}+\alpha\bar{q})+\bar{\mu}_b\lambda_b\alpha q,$$

$$F = \bar{\mu}_b\bar{\lambda}_b\alpha q+\mu_b\bar{\lambda}_b\alpha+\mu_b\lambda_b\alpha q,$$

$$G = \mu_b\bar{\lambda}_b\bar{\alpha}r_1(\lambda_b\alpha q+\bar{\lambda}_b\alpha)+\mu_b\bar{\lambda}_b\lambda_b(\bar{\alpha}+\alpha\bar{q})\bar{\alpha}.$$

Proof. Notice the structure of A_0, A_1 and A_2, assume $R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$, where R_{11}, R_{12} and R_{22} are all 2×2 matrices. Taking R into Eq. (1), have

$$
\begin{cases}
R_{11} = {R_{11}}^2 \begin{pmatrix} 0 & \lambda_v \alpha q \bar{\theta} + \bar{\lambda}_v \alpha\theta \\ 0 & \bar{\mu}_v \bar{\lambda}_v \alpha q \bar{\theta} \end{pmatrix} + R_{11} \begin{pmatrix} \bar{\lambda}_v \bar{\alpha} \bar{\theta} & \lambda_v \alpha \bar{q} \bar{\theta} + \lambda_v \bar{\alpha} \bar{\theta} \\ 0 & \bar{\mu}_v \bar{\lambda}_v (\bar{\alpha} + \alpha \bar{q}) \bar{\theta} + \bar{\mu}_v \lambda_v \alpha q \bar{\theta} \end{pmatrix} \\
+ \begin{pmatrix} 0 & 0 \\ 0 & \bar{\mu}_v \lambda_v (\bar{\alpha} + \alpha \bar{q}) \bar{\theta} \end{pmatrix} \\
R_{12} = {R_{11}}^2 \begin{pmatrix} 0 & \lambda_v \alpha q\theta + \bar{\lambda}_v \alpha\theta \\ 0 & \bar{\mu}_v \bar{\lambda}_v \alpha q\theta + \mu_v \bar{\lambda}_v \alpha + \mu_v \lambda_v \alpha q \end{pmatrix} \\
+ (R_{11}R_{12} + R_{12}R_{22}) \begin{pmatrix} 0 & \lambda_b \alpha q + \bar{\lambda}_b \alpha \\ 0 & \bar{\mu}_b \bar{\lambda}_b \alpha q + \mu_b \bar{\lambda}_b \alpha + \mu_b \lambda_b \alpha q \end{pmatrix} \\
+ R_{11} \begin{pmatrix} \bar{\lambda}_v \bar{\alpha} \theta & \lambda_v \bar{\alpha} \theta + \lambda_v \alpha \bar{q} \theta \\ \mu_v \bar{\lambda}_v \bar{\alpha} & (\mu_v \lambda_v + \bar{\mu}_v \bar{\lambda}_v)(\bar{\alpha} + \alpha \bar{q}) + \bar{\mu}_v \lambda_v \alpha q\theta \end{pmatrix} \\
+ R_{12} \begin{pmatrix} \bar{\lambda}_b \bar{\alpha} & \lambda_b (\bar{\alpha} + \alpha \bar{q}) \\ \mu_b \bar{\lambda}_b \bar{\alpha} & (\mu_b \lambda_b + \bar{\mu}_b \bar{\lambda}_b)(\bar{\alpha} + \alpha \bar{q}) + \bar{\mu}_b \lambda_b \alpha q \end{pmatrix} \\
+ \begin{pmatrix} 0 & 0 \\ 0 & \bar{\mu}_v \lambda_v (\bar{\alpha} + \alpha \bar{q})\theta \end{pmatrix} \\
R_{22} = {R_{22}}^2 \begin{pmatrix} 0 & \lambda_b \alpha q + \bar{\lambda}_b \alpha \\ 0 & \bar{\mu}_b \bar{\lambda}_b \alpha q + \mu_b \bar{\lambda}_b \alpha + \mu_b \lambda_b \alpha q \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \bar{\mu}_b \lambda_b (\bar{\alpha} + \alpha \bar{q}) \end{pmatrix} \\
+ R_{22} \begin{pmatrix} \bar{\lambda}_b \bar{\alpha} & \lambda_b (\bar{\alpha} + \alpha \bar{q}) \\ \mu_b \bar{\lambda}_b \bar{\alpha} & (\mu_b \lambda_b + \bar{\mu}_b \bar{\lambda}_b)(\bar{\alpha} + \alpha \bar{q}) + \bar{\mu}_b \lambda_b \alpha q \end{pmatrix}
\end{cases}
$$

From the first and third equations, get$R_{11} = \begin{pmatrix} 0 & 0 \\ 0 & r_1 \end{pmatrix}$ and $R_{22} = \begin{pmatrix} 0 & 0 \\ r_4 & r_5 \end{pmatrix}$ respectively. Then, taking R_{11} and R_{22} into the second equation, finally obtain $R_{12} = \begin{pmatrix} 0 & 0 \\ r_2 & r_3 \end{pmatrix}$ are as shown in Theorem 2.

3.3 Stationary Distribution and Performance Measures

If $\bar{\mu}_b(\bar{\lambda}_b \bar{\alpha} - 1)(\alpha q - \lambda_b) < \mu_b(\bar{\lambda}_b + \lambda_b q)\alpha$, Let $\{Q, J\}$ be the stationary limit of the $\{Q_n, J_n\}$. And denote

$$
\begin{aligned}
\pi_k &= (\pi_{k0}, \pi_{k1}, \pi_{k2}, \pi_{k3}), k \geq 0, \\
\pi_{kj} &= P\{Q = k, J = j\} = \lim_{n \to \infty} P\{Q_n = k, J_n = j\}, (k, j) \in \Omega.
\end{aligned}
$$

Note that, $\pi_{02} = 0$.

Theorem 3. *If* $\bar{\mu}_b(\bar{\lambda}_b \bar{\alpha} - 1)(\alpha q - \lambda_b) < \mu_b(\bar{\lambda}_b + \lambda_b q)\alpha$, *the stationary distribution of* Q, J *is given by*

$$\begin{cases} \pi_{k0} = 0, & k \geq 1 \\ \pi_{k1} = \pi_{01} {r_1}^k, & k \geq 0 \\ \pi_{k2} = \pi_{01}(r_2 {r_1}^{k-1} + \frac{r_3 r_4}{r_5 - r_1}({r_5}^{k-1} - {r_1}^{k-1})) + \pi_{03} r_4 {r_5}^{k-1}, & k \geq 1 \\ \pi_{k3} = \pi_{01} \frac{r_3}{r_5 - r_1}({r_5}^k - {r_1}^k) + \pi_{03} {r_5}^k, & k \geq 0, \end{cases} \tag{2}$$

where

$$\pi_{00} = \left[1 + (\frac{1 + r_2}{1 - r_1} + \frac{r_3(1 + r_4)}{(1 - r_1)(1 - r_5)})Y + \frac{1 + r_4}{1 - r_5} U\right]^{-1},$$

$$Y = \frac{\lambda_v \bar{\theta}}{1 - \bar{\mu}_v \bar{\lambda}_v \bar{\theta} - \bar{\mu}_v \bar{\lambda}_v \alpha q \bar{\theta} r_1}, U = \frac{\lambda_v(1 - \bar{\lambda}_v \bar{\theta}) - \bar{\mu}_v \bar{\lambda}_v \lambda_v \alpha q \bar{\theta} r_1}{\mu_b \bar{\lambda}_b(1 - \bar{\mu}_v \bar{\lambda}_v \bar{\theta} - \bar{\mu}_v \bar{\lambda}_v \alpha q \bar{\theta} r_1)},$$

$$\pi_{01} = \frac{\lambda_v \bar{\theta}}{1 - \bar{\mu}_v \bar{\lambda}_v \bar{\theta} - \bar{\mu}_v \bar{\lambda}_v \alpha q \bar{\theta} r_1} \pi_{00}, \pi_{03} = \frac{\lambda_v(1 - \bar{\lambda}_v \bar{\theta}) - \bar{\mu}_v \bar{\lambda}_v \lambda_v \alpha q \bar{\theta} r_1}{\mu_b \bar{\lambda}_b(1 - \bar{\mu}_v \bar{\lambda}_v \bar{\theta} - \bar{\mu}_v \bar{\lambda}_v \alpha q \bar{\theta} r_1)} \pi_{00}.$$

Proof. Using the matrix-analytic method [1], have

$$\pi_k = (\pi_{k0}, \pi_{k1}, \pi_{k2}, \pi_{k3}) = \pi_0 R^k = (\pi_{00}, \pi_{01}, 0, \pi_{03}) R^k, k \geq 0. \tag{3}$$

Taking

$$R^k = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & {r_1}^k & r_2 {r_1}^{k-1} + \frac{r_3 r_4}{r_5 - r_1}({r_5}^{k-1} - {r_1}^{k-1}) & \frac{r_3}{r_5 - r_1}({r_5}^k - {r_1}^k) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & r_4 {r_5}^{k-1} & {r_5}^k \end{pmatrix}, k \geq 1$$

into Eq. (3), then Eq. (2) is obtained. Meanwhile, π_0 satisfies the following two equations

$$\pi_0(B_1 + R A_2) = \pi_0, \tag{4}$$

$$\pi_0(I - R)^{-1} e = 1. \tag{5}$$

From Eq. (4), get π_{01} and π_{03} as shown in Theorem 3. Then, taking $\pi_{01}, \pi_{02} = 0$ and π_{03} into Eq. (5), obtain π_{00} as shown in Theorem 3. So far, Theorem 3 is proved.

Using the stationary distribution derived above, some performance measures can be obtained as follows.

(1) The probability that the server is busy

$$P_b = \sum_{k=0}^{\infty} \pi_{k1} + \sum_{k=0}^{\infty} \pi_{k3} = \frac{1 - r_5 + r_3}{(1 - r_1)(1 - r_5)} \pi_{01} + \frac{1}{1 - r_5} \pi_{03}.$$

(2) The probability that the server is free

$$P_f = \sum_{k=0}^{\infty} \pi_{k0} + \sum_{k=1}^{\infty} \pi_{k2} = \pi_{00} + \frac{r_2 - r_2 r_5 + r_3 r_4}{(1-r_1)(1-r_5)}\pi_{01} + \frac{r_4}{1-r_5}\pi_{03}.$$

(3) The probability that the server is in a regular busy period

$$P_r = \sum_{k=0}^{\infty} \pi_{k2} + \sum_{k=0}^{\infty} \pi_{k3} = \frac{r_2 + r_3 - r_2 r_5 + r_3 r_4}{(1-r_1)(1-r_5)}\pi_{01} + \frac{1+r_4}{1-r_5}\pi_{03}.$$

(4) The probability that the server is in a regular busy period

$$P_v = \sum_{k=0}^{\infty} \pi_{k0} + \sum_{k=0}^{\infty} \pi_{k1} = \pi_{00} + \frac{1}{1-r_1}\pi_{01}.$$

(5) The mean number of customers in the orbit

$$\begin{aligned} E[L] &= \sum_{k=1}^{\infty} k(\pi_{k1} + \pi_{k2} + \pi_{k3}) \\ &= \pi_{01}\frac{(r_1+r_2)(1-r_5)^2 + r_3 r_4(2-r_1-r_5) + r_3(1-r_1 r_5)}{(1-r_1)^2(1-r_5)^2} + \pi_{03}\frac{r_4+r_5}{(1-r_5)^2}. \end{aligned}$$

(6) The mean number of customers in the system

$$E[\widetilde{L}] = \sum_{k=1}^{\infty} k\pi_{k2} + \sum_{k=0}^{\infty} (k+1)(\pi_{k1} + \pi_{k3}) = E[L] + P_b.$$

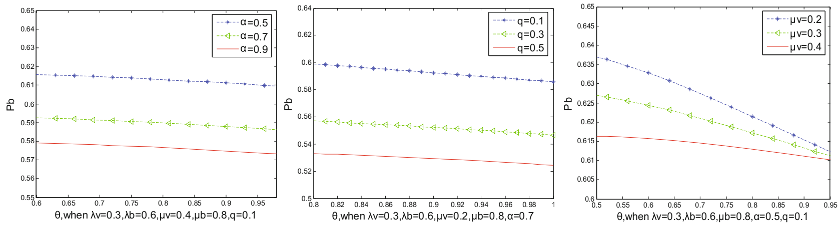
3.4 Empirical Research

This section provides some numerical examples to illustrate the effect of varying parameters on the performance measures of the system. Assume λ_v equals 0.3, λ_b equals 0.6 and μ_b equals 0.8. Under the condition that $\bar{\mu}_b(\bar{\lambda}_b\bar{\alpha} - 1)(\alpha q - \lambda_b) < \mu_b(\bar{\lambda}_b + \lambda_b q)\alpha$, plot Figs. 3, 4, 5, 6, 7, 8, 9 and 10, which illustrate the effect of vacation period rate θ, retrial rate α, departure rate q and service rate μ_v on the performance measures P_b and $E[L]$.

Remark 1. Considering the constraint of the stationary condition and the display of the graph, the values of the parameters are the most suitable.

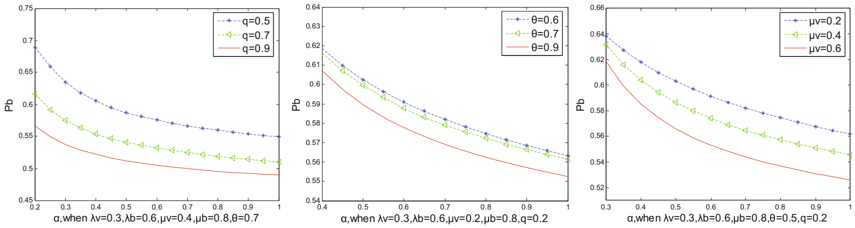
(1) The numerical analysis of P_b
Figures 4, 5, 6 and 7 illustrate the effect of parameters on the probability that the server is busy. The effect and the reasons which cause the effect are discussed.

Figure 4 illustrates the effect of θ on the probability that the server is busy. Figure 4(a) is for different α, Fig. 4(b) is for different q and Fig. 4(c) is for different μ_v. Obviously, whether in Fig. 4(a), (b) or (c), the probability P_b decreases



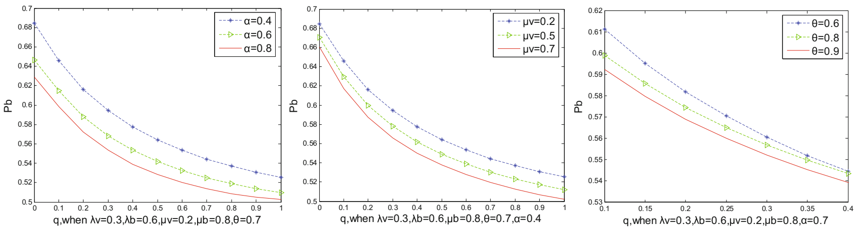
(a) (b) (c)

Fig. 4. (a) The probability that the server is busy versus θ for different α (b) The probability that the server is busy versus θ for different q (c) The probability that the server is busy versus θ for different μ_v



(d) (e) (f)

Fig. 5. (d) The probability that the server is busy versus α for different q (e) The probability that the server is busy versus α for different θ (f) The probability that the server is busy versus α for different μ_v

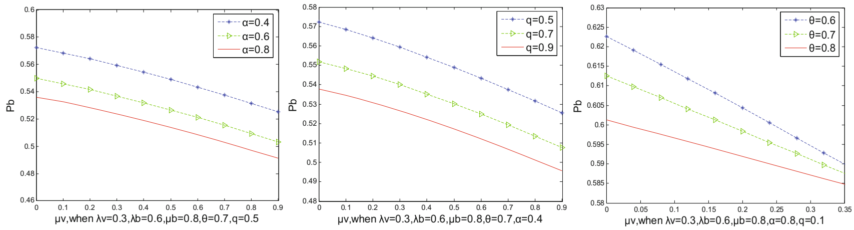


(g) (h) (i)

Fig. 6. (g) The probability that the server is busy versus q for different α (h) The probability that the server is busy versus q for different μ_v (i) The probability that the server is busy versus q for different θ

gradually with the rate θ increasing regardless of the change of α, q and μ_v. What causes this trend is the fact that the average service rate becomes smaller with θ increasing. Note that the effect of θ on P_b becomes weaker with μ_v increasing. Meanwhile, because $P_b + P_f = 1$, the probability that the sever is free P_f should show the opposite trend of change (the same below).

Figure 5 illustrates the effect of α on the probability that the server is busy. Figure 5(d) is for different q, Fig. 5(e) is for different θ and Fig. 5(f) is for



(j) (k) (l)

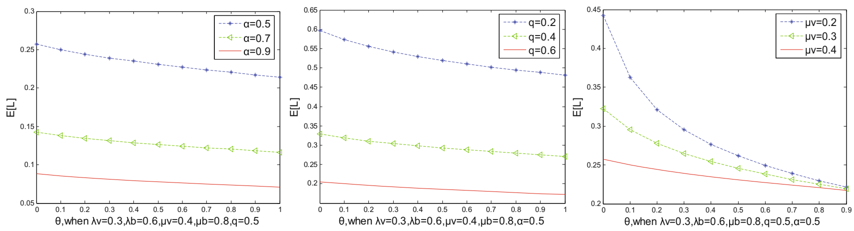
Fig. 7. (j) The probability that the server is busy versus μ_v for different α (k) The probability that the server is busy versus μ_v for different q (l) The probability that the server is busy versus μ_v for different θ

different μ_v. Among these three figures, P_b decreases with the rate α increasing regardless of the change of α, q and μ_v. This is because the number of the customers who give up receiving service increases with μ_v increasing.

Figure 6 illustrates the effect of q on the probability that the server is busy. Figure 6(g) is for different α, Fig. 6(h) is for different μ_v and Fig. 6(i) is for different θ. Obviously, P_b decreases with the rate q increasing among these three figures. This is because the number of the customers who leave the system due to the failure of retrial increases with q increasing.

Figure 7 illustrates the effect of v on the probability that the server is busy. Figure 7(j) is for different α, Fig. 7(k) is for different q and Fig. 7(l) is for different θ. Obviously, P_b decreases with the rate μ_v increasing. This is because the average service rate increases with μ_v increasing. From the change of the other three parameters, this trend is consistent. Besides, the effect of μ_v on P_b also becomes weaker with θ increasing.

(2) The numerical analysis of $E[L]$
Figures 8, 9, 10 and 11 illustrate the effect of parameters on the mean number of customers in the orbit. The effect and the reasons that cause the effect are also discussed.



(m) (n) (o)

Fig. 8. (m) The mean number of customers in the orbit versus θ for different α (n) The mean number of customers in the orbit versus θ for different q (o) The mean number of customers in the orbit versus θ for different μ_v

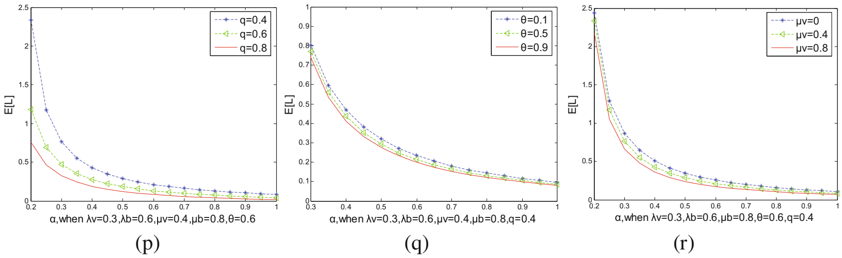


Fig. 9. (p) The mean number of customers in the orbit versus α for different q (q) The mean number of customers in the orbit versus α for different θ (r) The mean number of customers in the orbit versus α for different μ_v

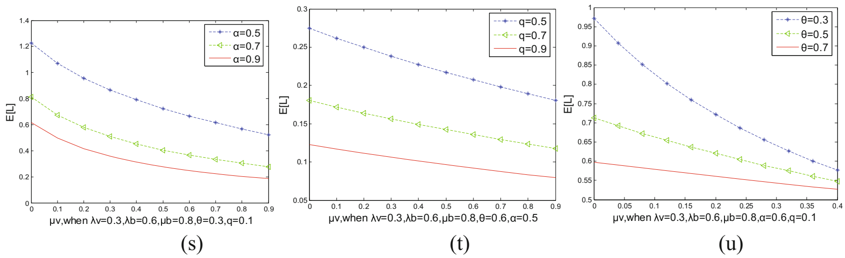
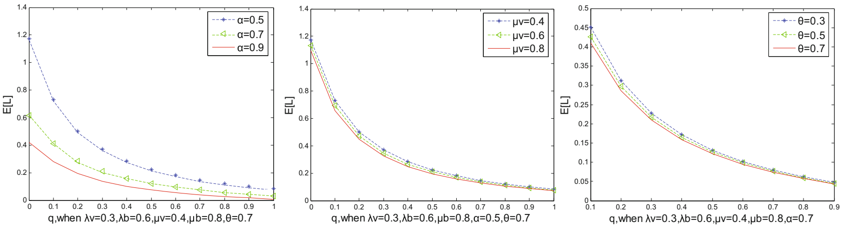


Fig. 10. (s) The mean number of customers in the orbit versus μ_v for different α (t) The mean number of customers in the orbit versus μ_v for different q (u) The mean number of customers in the orbit versus μ_v for different θ

Figure 8 illustrates the effect of θ on the mean number of customers in the orbit. Figure 8(m) is for different α, Fig. 8(n) is for different q and Fig. 8(o) is for different μ_v. With the rate θ increasing, the time during which the sever is in a working vacation period is reduced, then the average service decreases, which leads to the decreasing trend of the mean number of customers $E[L]$ as shown in the figures. It is worth noting that the effect of θ on $E[L]$ becomes weaker with μ_v increasing, which is consistent with the effect of θ on P_b in the Fig. 4(c).

Figure 9 illustrates the effect of α on the mean number of customers in the orbit. Figure 9(p) is for different q, Fig. 9(q) is for different θ, Fig. 9(r) is for different μ_v. For the same q, θ and μ_v, $E[L]$ decreases with α increasing. It is because the number of the customers who leave the orbit due to success or failure of retrial increase with the retrial rate α increasing. Furthermore, with α increasing, the effect of it on $E[L]$ becomes weaker.

Figure 10 illustrates the effect of μ_v on the mean number of customers in the orbit. Figure 10(s) is for different α, Fig. 10(t) is for different $q$and Fig. 10(u) is for different θ. Obviously, no matter how α, q and θ change, $E[L]$ decrease with μ_v increasing. This is because the average service rate increases with μ_v increasing. From Fig. 10(u), we can see that the effect of μ_v on $E[L]$ becomes



(v) (w) (x)

Fig. 11. (v) The mean number of customers in the orbit versus q for different α (w) The mean number of customers in the orbit versus q for different μ_v (x) The mean number of customers in the orbit versus q for different θ

smaller with θ increasing, which is consistent with the effect of μ_v on P_b in the Fig. 7(l).

Figure 11 illustrates the effect of q on the mean number of customers in the orbit. Figure 11(v) is for different α, Fig. 11(w) is for different v and Fig. 11(x) is for different θ. With q increasing, the number of customers who leave the system due to failure in retrial increases, which leads to the decreasing trend of $E[L]$ as shown in the figures. Besides, the effect of q on $E[L]$ becomes weaker with q increasing. And the lager α is, the smaller the effect of q on $E[L]$ becomes, which means that the departure rate of customers will diminish the impact on the mean number of customers in the orbit when the load of the orbit becomes lighter.

(3) Performance optimization

Using the performance measures derived in Sect. 3.3, this section designs a cost function, i.e., the average total cost of the system per unit time, and establishes an optimization model on the basis of the cost function. Then we use the optimization model to optimize the queueing model established in Sect. 3.1 under the environment of web service system, and obtain the optimal parameters and the minimum average total cost per unit time.

To design the cost function, the cost structure is defined firstly. Assume there are two costs in the system, one is the waiting cost, and the other is the service cost. Assume the average waiting cost of each customer per unit time is c, the average service cost of each server in a regular busy period per unit time is a_r, and the average service cost of each server in a working vacation period per unit time is a_v Then, the average total cost of the system per unit time M can be expressed as

$$M = cE[\widetilde{L}] + a_r P_r + a_v P_v, \tag{6}$$

i.e., cost function. Based on Eq. (6), establish an optimization model as

$$
\min = cE[\widetilde{L}] + a_r P_r + a_v P_v
$$
$$
s.t. \begin{cases} \bar{\mu}_b(\bar{\lambda}_b \, \bar{\alpha} - 1)(\alpha q - \lambda_b) < \mu_b(\bar{\lambda}_b + \lambda_b q)\alpha \\ E[\widetilde{L}] \geq 0 \\ 0 < P_r, P_v < 1 \\ 0 < \mu_v < \mu_b < 1, 0 < \theta, \alpha < 1. \end{cases}
$$

Then, consider using the optimization model above to optimize the Geo/Geo/1 retrial queue with working vacation interruption and nonpersistent customers established in Sect. 3 under the environment of web service system. Assume the average waiting cost of each customer per unit time is 5, the average service cost of each server in a regular busy period per unit time is 10, the average service cost of each server in a working vacation period per unit time is 6, the probability that the user requests service during a regular busy period is 0.6, the probability that the user requests service during a working vacation period is 0.3, the retrial rate is 0.6 and the probability that the user gives up receiving service due to the failure of retrial is 0.2. Taking these parameters into the optimization model and using LINGO to solve it, get $\mu_v = 0.277, \mu_b = 0.739, \theta = 0.125, M = 11.359$. The result means that the average total cost of the system per unit time is 11.359, i.e., the minimum sum of the average waiting cost and the average service cost per unit time, now, the service rate in a regular busy period and the service rate in a working vacation period are 0.739, 0.277 respectively, and the vacation period rate θ is 0.125. That is to say, in order that the performance of the web service system is the best, managers should adjust the actual service capabilities of the server to guarantee that the service time in a regular busy period follows a geometric distribution with parameter 0.739, the service time in a working vacation period follows a geometric distribution with parameter 0.277 and the working vacation time follows a geometric distribution with parameter 0.125.

4 Conclusions

With the development of the internet, the web service industry has put forward many new requirements to the design of queueing model. This paper provides a queueing solution for web service. After presenting three practical problems that web service encounters in real life, a queuing model which can deal with these problems is constructed and a series of necessary researches are carried out on the model. Then, based on the results of the researches, the numerical analysis and model optimization are conducted, which provide some valuable consultations for the web service management under the corresponding circumstances. In addition, since the model proposed in this paper is a single server, which is somehow discrepant with the reality of multi-server web service, the number of services can be pending in the future research to better cope with the actual situation.

References

1. Burr T (2001) Introduction to matrix analytic methods in stochastic modeling. Technometrics 95(3):1379
2. Dimitriou I (2015) A retrial queue for modeling fault-tolerant systems with checkpointing and rollback recovery. Comput Ind Eng 79:156–167
3. Kobayashi H, Konheim AG (1977) Queueing models for computer communications system analysis. IEEE Trans Commun 25(1):2–29
4. Kosten L (1947) On the influence of repeated calls in the theory of probabilities of blocking. De Ingenieur 49(1):947
5. Li J, Tian N (2007) The discrete-time GI/Geo/1 queue with working vacations and vacation interruption. Appl Math Comput 185(1):1–10
6. Li J, Tian N (2007) The M/M/1 queue with working vacations and vacation interruptions. J Syst Sci Syst Eng 16(1):121–127
7. Li J, Tian N (2008) Analysis of the discrete time Geo/Geo/1 queue with single working vacation. Qual Technol Quant Manage 5(1):77–89
8. Li T, Wang Z, Liu Z (2012) Geo/Geo/1 retrial queue with working vacations and vacation interruption. J Appl Math Comput 39(1):131–143
9. Liu Z, Song Y (2013) Geo/Geo/1 retrial queue with non-persistent customers and working vacations. J Appl Math Comput 42(1):103–115
10. Meisling T (1958) Discrete-time queuing theory. Oper Res 6(1):96–105
11. Palm C (1953) Methods of judging the annoyance caused by congestion. Tele 4(189208):4–5
12. Phung-Duc T (2014) Multiserver retrial queues with two types of nonpersistent customers. Asia Pac J Oper Res 31(2):144–171
13. Rajadurai P (2015) Analysis of M [X]/G/1 retrial queue with two phase service under bernoulli vacation schedule and random breakdown. In J Math Oper Res 7(1):19–41
14. Servi LD, Finn SG (2002) M/M/1 queues with working vacations (M/M/1/WV). Perform Eval 50(1):41–52
15. Tao L, Zhang L, Xu X (2013) An M/G/1 queue with single working vacation and vacation interruption under Bernoulli schedule. Appl Math Model 37(3):1564–1579
16. Tian N, Zhang Z (2006) Vacation queueing models: theory and applications. Springer, New York
17. Tian N, Ma Z, Liu M (2008) The discrete time Geom/Geom/1 queue with multiple working vacations. Appl Math Model 32(12):2941–2953
18. Yang T, Li H (1995) On the steady-state queue size distribution of the discrete-time Geo/G/1 queue with repeated customers. Queueing Syst 21(1):199–215

Pan-Tourism Urbanization Model Based on System Dynamics: A Case Study of Barkam

Lin Hu, Lu Gan(✉), and Qian Yang

College of Architecture and Urban-Rural Planning, Sichuan Agricultural University, Dujiangyan 611830, People's Republic of China
ganlu_soarpb@sicau.edu.cn

Abstract. With the rapid development of urbanization, tourism-driven urbanization process has become an important development strategy in China. There are many examples of the tourism urbanization, however, the integration with industries is scarce. In order to describe the characterization of the industrial clusters system, this paper uses SPSS [8] to analyze the data. And it also uses DYNAMO simulation language and Vensim [2] to establish the pan-tourism urbanization model based on the theory of industrial cluster and system dynamics. Simultaneously, taking Barkam city in the northwestern Tibetan area of Sichuan as an example, this paper analyzes the operation mechanism of pan-tourism urbanization mode to find the optimal system structure, which can provide corresponding scientific basis and theoretical support for promoting the development of urban and rural areas, improving the level of new urbanization and building new towns.

Keywords: System dynamics · Pan-tourism · Industrial cluster · Simulation

1 Introduction

In recent years, the model of new urbanization has been widely concerned. Tourism industry and related industries develop interactively in order to promote the new urbanization and the coordinated development of new rural construction. Therefore, the model of pan-tourism industry integration and industrial cluster development is an effective way to promote the development of urbanization, which is proposed by experts and scholars at home and abroad.

Some scholars have made some research ideas. Alexandrova [1] aims to review tourism cluster initiatives in Russia, particularly, tourism cluster formation processes. Jackson [6] takes four towns in Australia as an example, exploring the effectiveness of correlation factors of these four regions, and studying the competitive advantage of cluster theory. However, they only concentrate on the research of the existing tourism industrial clusters evaluation and focus on using descriptive language to assess or predict which lacks a practical and objective evaluation system.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_35

Jackson [5] probes into the development of regional tourism clusters, and he finds the correctness in theory of the Porter's [12] model. Tourism cluster may become a way to convert comparative advantage into competitive advantage and make better use of existing tourism resources in China. It can be more adapted to the characteristics of Chinese region. Qian [13] questions Mullins's [10] classical theory of tourism urbanization, and he shows the idea that the process of urbanization driven by modern tourism is not the product of post-modernization of urban cultural expression, however, it is the product of tourism-related services based on relatively standardized regulations. The two scholars focus on the comparative analysis of relevant theoretical models and existing data. The research model they put forward has positive effects on the development of domestic tourism, but they ignore the systematic description of tourism urbanization.

This paper is organized as follows. In Sect. 2, the pan-tourism urbanization is introduced and then we proposed a differential pan-tourism system model with industrial clusters theory. In Sect. 3, on the basis of pan-tourism system with the actual situation in Barkam, we investigate the relevant data, including the number of tourists, tourism revenue and other data about urbanization. Once we have carried out regression analysis, we can get simulation results for pan-tourism urbanization system. And we analyze the results and give recommendations. In Sect. 4, we make a conclusion about the effectiveness of such a model and discuss the shortcomings of this model.

2 Modeling

2.1 Concept Model

System dynamics is a science of dynamic complexity of the system developed by the Professor Forester of MIT, which emphasizes the overall consideration of the system and the interaction of each composition of the system. It is designed to deal with the problems of high order, nonlinear and multiple feedbacks in complex system [17]. Based on the method of qualitative and quantitative analysis and the system dynamic simulation experiments, which can forecast the dynamic change, the decision makers can observe the simulation results and take different measures in various contexts.

In this paper, the process of system dynamics modeling can be divided as follows: firstly, according to the characteristics of problems, we realize the purpose of modeling and determine the system boundaries to avoid the system architecture which will be too large or too small; secondly, we analyze the system structure to divide the system hierarchy and sub modules, determine the causal feedback loop and define the interaction mechanism among elements; thirdly, through the regression analyzing of the data, we establish the system dynamic mathematical model completely; finally, we analyze the simulation results which are obtained from the model running and evaluate the model in order to identify the problems and improve the model.

Based on the feedback control theory and the concept of industrial cluster development, this paper constructs the dynamic analysis model of pan-tourism

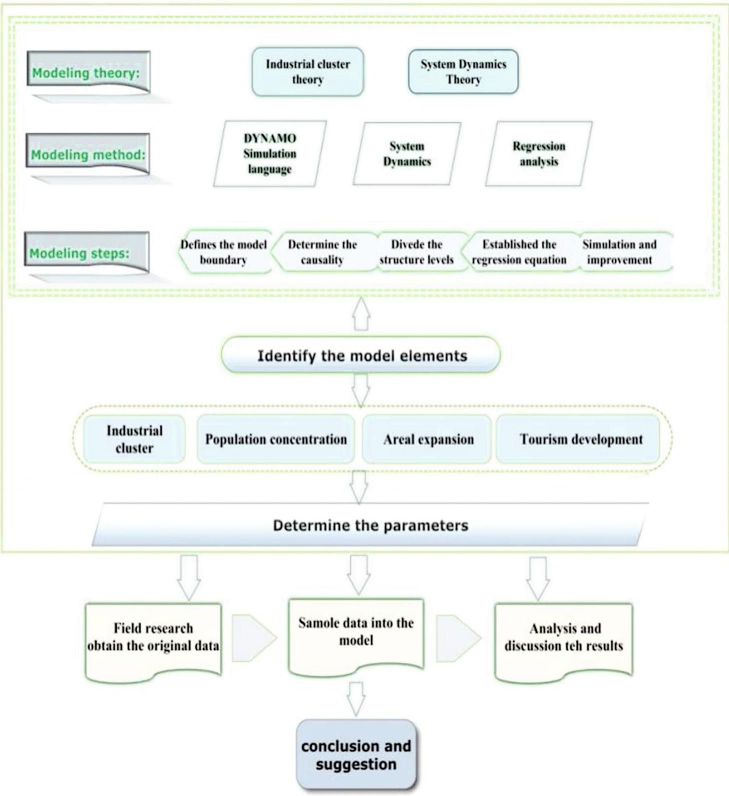


Fig. 1. Flow chart for research process

urbanization system to find out the basic reasons, key factors, changing rules and internal relations further. The research ideas are shown in Fig. 1.

2.2 System Elements Determination

System dynamics believes that the fundamental power and the direct cause of the whole socio-economic development come from the inner system, which determines the behavior pattern and characteristics. Besides, the system boundaries define the minimum number of units. Only defining the boundaries can the system solve structural problems [15]. Traditional tourism is only a single industry with simple form, while the pan-tourism urbanization model is the integration of pan-tourism industry development.

Pan-tourism industrial clusters have three main modes, such as the feedback, promotion and interaction modes [4]. The first one is the feedback model, referring to the non-profit service industries, which provide the tourism industry with resourceful and environmental infrastructures. They receive economic benefits from tourism and make themselves developing stably and healthily with tourism. The second promotion model means industries which cannot integrate with the tourism to produce new products or formats, but their development will be promoted when tourism developed vigorously. For instance, the lodging, catering and transportation industry can supply support elements for the tourism while the wholesale and retail trade, real estate, warehousing postal services, construction and industry which offering corollary facility to tourism. The third one is the interaction mode, referring to integration of agriculture and tourism, which can form a new tourism products and formats, rich tourism content and enhance the value of tourism products. Not only further does the mode develop the new tourism market, but also it promotes the industry and tourism industry to develop collaboratively.

The development of tourism industry leads to the development of related industries and forms pan-tourism industrial clusters, therefore it brings about the agglomeration of population and elements, and promotes the process of urbanization in local area. Urbanization is the process of the evolution of the economic structure, mode of production, life style and social concept after rural population and various factors constantly gathered to the city. The level of urbanization is usually measured by the proportion of urban population to total population.

Actually, urbanization is not just a regional change in the nature of the population. The connotation of urbanization is very abundant which means a regional change process based on the non-agricultural population and non - agricultural economy [16], and it's overgeneralization if the urbanization level only measured by the population ratio. This paper references to Fang [3], Sun [14] and other scholar's defining of urbanization degree and combines with the development characteristics of pan-tourist urbanization to define urbanization degree of pan-tourism urbanization system by using two indicators of urbanization and economic urbanization.

The population urbanization rate is usually used to measure the level of urbanization, and it is an important symbol to measure a country's level of economic development [7]. The integration development of tourism and related industries influence the changes of the industrial structure, reduce the employees of the primary industry continuously and increase the employees of the second and tertiary Industry. In addition, the urban employment increases through the transfer of labor, while gradually realizing the transformation of farmers' identity, then completing the population in situ urbanization.

The economic urbanization rate reflects the non-agricultural process of the overall economic structure, characterized by the proportion of the second and third industries gradually increased while proportion of primary industry gradually decreased. Tourism drives pan-tourism industrial clusters, whose output values are calculated from the second and third industrial output values, location entropies and the contribution rates [18]. The economic urbanization rates are the proportion of cluster output values to GDPs.

2.3 System Causality Analysis

Pan-tourism urbanization system includes three subsystems: tourism subsystem, industrial cluster subsystem and urbanization subsystem. This paper chooses the relevant variables from the perspective of pan-tourism industrial cluster and urbanization, and analyzes the system elements from the perspective of tourism driven urbanization.

The pan-tourism industrial clusters and urbanization interactive development system consist of three basic feedback loops: firstly, tourism drives pan-tourism cluster through industrial integration, increasing the cluster output values and improving the level of urbanization; secondly, industrial cluster increases employment in the second and third industries. The transfer of the labor also increases the urban population and raises the urbanization level; thirdly, the development of region economy and the increase of urban population promote urban construction, thus enhancing tourism reception capacity and developing the tourism industry. The causal relationship is shown in Fig. 2.

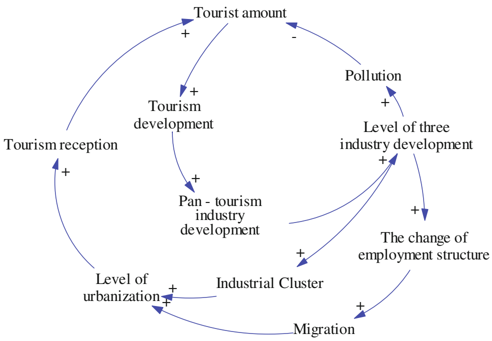


Fig. 2. Pan-tourism urbanization causality graph

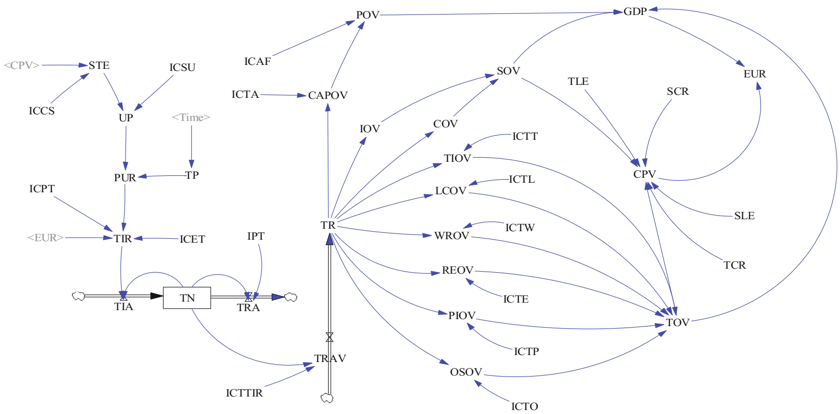


Fig. 3. Pan-tourism urbanization system flow diagram

Table 1. Meaning of the main variables and influence factors

Notation	Variable name	Unit
TN	Number of tourists	10^4 persons
TR	Tourism revenue	10^4 yuan
TIR	Tourist increase rate	No dimension
TIA	Tourist increase amount	10^4 persons
TRA	Tourist reduction amount	10^4 persons
TRAV	Tourism revenue added value	10^4 yuan
PIOV	The primary industry output value	10^4 yuan
SOV	The secondary industry output value	10^4 yuan
TOV	The tertiary industry output value	10^4 yuan
SLE	Secondary industry location entropy	No dimension
TLE	Tertiary industry location entropy	No dimension
SCR	Secondary industry contribution rate	No dimension
TCR	Tertiary Industry contribution rate	No dimension
EUR	Economic urbanization rate	No dimension
CPV	Cluster production value	10^4 yuan
STE	Employment in the second and third industries	10^4 persons
UP	Urban population	10^4 persons
TP	Total population	10^4 persons
PUR	Population urbanization rate	No dimension
IPT	Impact index of pollution on tourists	No dimension
CAPOV	Characteristic agricultural products output value	10^4 yuan
IOV	Industrial output value	10^4 yuan
COV	Construction industry output value	10^4 yuan
TIOV	Transportation industry output value	10^4 yuan
LCOV	Lodging and catering industry output value	10^4 yuan
WROV	Wholesale and retail output value	10^4 yuan
REOV	Real estate output value	10^4 yuan
PIOV	Postal industry output value	10^4 yuan
OSOV	Other service industry output value	10^4 yuan
ICAF	Influence coefficient of CAPOV on POV	No dimension
ICTA	Influence coefficient of TR on CAPOV	No dimension
ICTI	Influence coefficient of TR on IOV	No dimension
ICTC	Influence coefficient of TR on COV	No dimension
ICTT	Influence coefficient of TR on TOV	No dimension
ICTL	Influence coefficient of TR on LCOV	No dimension
ICTW	Influence coefficient of TR on WROV	No dimension
ICTE	Influence coefficient of TR on REOV	No dimension
ICTP	Influence coefficient of TR on PIOV	No dimension
ICTO	Influence coefficient of TR on OSOV	No dimension
ICTT	Influence Coefficient of TN on TRAV	No dimension
ICET	Influence coefficient of EUR on TRAV	No dimension
ICPT	Influence coefficient of PUR on TN	No dimension
ICCS	Influence Coefficient of CPV on STE	No dimension
ICSU	Influence Coefficient of STE on UP	No dimension

2.4 System Flow Diagram and Parameter Equation

(1) System Flow Diagram

As shown in Fig. 3, two flow variables, three rate variables and multiple auxiliary variables are selected to construct the system flow diagram based on the causal feedback map. The meaning of the main variables and influence factors in this model is shown in Table 1.

(2) Index System and Parametric Equation

The parametric equations are determined according to the logical relationship between system variables. There are two level variables named number of tourists and tourism revenue, and the three rate variables are tourist increase amount, tourist reduction amount and tourism revenue added value. The auxiliary variables are logically related to the level and rate variables. The constants include the influence coefficients, the contribution rates, the location entropy, and the total population expressed by the table function. The main parameters and equations are shown in Table 2.

Table 2. Main variables and equations

Variable type	Variable name	Equation
Level variable	TN	INTEG (TIA − TRA, Initial value)
	TR	INTEG (TRAV, Initial value)
Rate variable	TIA	TN × TIR
	TRA	TN × IPT
	TRAV	ICTTIR × TN
Auxiliary variable	TIR	ICETEUR+ICPTPUR
	CAPOV	TR × ICTA
	IOV	ICTI × TR
	COV	ICTC × TR
	TIOV	ICTT × TR
	LCOV	ICTL × TR
	WROV	ICTW × TR
	REOV	ICTE × TR
	PIOV	ICTP × TR
	OSOV	ICTO × TR
	POV	ICAF × CAPOV
	SOV	IOV + COV
	TOV	Regression equation
	CPV	SLE × SOV × SCR + TLE × TOV × TCR
	GDP	POV + SOV + TOV
	EUR	CPV/GDP
	TEP	ICCS × Ln(CPV)
	UP	Regression equation
	PUR	UP/TP
Constant	SLE, TLE, SCR, TCR	No dimension
	TP	Table function

3 Case Analysis

3.1 Data Sources

In this paper, Barkam city in northwest Tibetan area of Sichuan is selected as the study object. The urbanization mode driven by tourism has been well reflected at here. Based on the actual situation of Barkam, the pan-tourism urbanization system is simulated [9]. The primary data is collected from Barkam's statistical yearbook and Sichuan Province statistical yearbook. Although there are some uncertain values, such as the total population and urban population, which are strongly affected by the social environment and natural disasters, it can be explained reasonably to make the model reflect the general trend of pan-tourism system.

The simulation step length of this model is set to 1 year, and the simulation year is from 2011 to 2025. The initial values for 2011 are determined based on the Barkam's statistical yearbook: the number of tourists was 505,800; tourism revenue is 434.48 million Yuan. Table 3 shows the values of parameters determined by regression analysis, table function and other methods.

Table 3. Parameter's value

Parameter	Value	Explanation
SLE	0.354	(Regional secondary or tertiary industry output/Regional GDP)/(national secondary or tertiary industry output/GDP)
TLE	1.654	
SCR	0.15	The contribution degree of urban secondary and tertiary industry to regional GDP
TCR	0.75	
ICET	0.172	Regression equation:
ICPT	4.85	$TIR = (0.172099 \times EUR - 4.85 \times PUR) + 2.4092$
ICT	0.964, 0.996, 0.486, 1.164, 2.094, 0.151	Regression equation: $TOV = 0.964 \times TIOV + 0.996 \times LCOV$ $+0.486 \times WROV + 1.164 \times REOV$ $+2.094 \times PIOV + 0.151 \times OSOV$
ICSU	0.42, −1.788	Regression equation: $UP = -1.788 \times STE + 0.42 \times STE^2 + 4.616$
ICCS, ICTTIR, ICTA,	0.514, 101.277, 0.225,	Regression analysis determines the parameters
ICTI, ICTC, ICTT,	0.298, 0.068, 0.384,	
ICTL, ICTW, ICTE,	0.316, 0.081, 0.154,	
ICTP, ICTO	1.662, 0.788	
IPT	0.04	The extent of the impact of pollution on the reduction of visitors
TP	Table function	With Lookup (([(2011,5.9)-(2015,5.9)],(2011,5.9),(2012,5.9),(2013,5.9),(2014,5.9),(2015,5.9))

3.2 Validity Test

(1) Historical Inspection

In this paper, the comparison of simulation results with actual data is conducted to test the effectiveness of the model. The simulated data, actual data and error rate for 2011–2015 are shown in Table 4. These data show that the error rate is kept within the 10% controllable range, which indicates the model can forecast the running state of the real system with good fitting degree.

(2) Operation inspection

The step length of these fitting values of tourism revenue is 1 year, and thus we select the step length of 0.5 years and 0.25 years to make comparison. It can be seen that curve of tourism revenue of Barkam changes implicitly under different step lengths of simulation. And the average error rate is less than 10%, which means the analogue system is stable (see Fig. 4).

3.3 Simulation Results

(1) Tourists and Tourism Revenue

Figures 5 and 6 show the simulation results of the tourists and the tourism revenue in Barkam City. The number of tourists of Barkam will increase steadily

Table 4. Historical data and simulation results analysis

Variable name	Value type	2011	2012	2013	2014	2015
$TN/10^4$ persons	Actual value	50.58	64.57	77.3	85.74	92.02
	Simulation value	50.58	60.8498	73.1725	86.4152	99.9342
	Error rate/%	0	−5.761	−5.34	0.787	8.601
$TR/10^4$ yuan	Actual value	43448	57725	69642	76572	84565
	Simulation value	43448	57023.2	64819.7	73864.2	84249.9
	Error rate/%	0	−1.216	−6.924	−3.536	−0.373

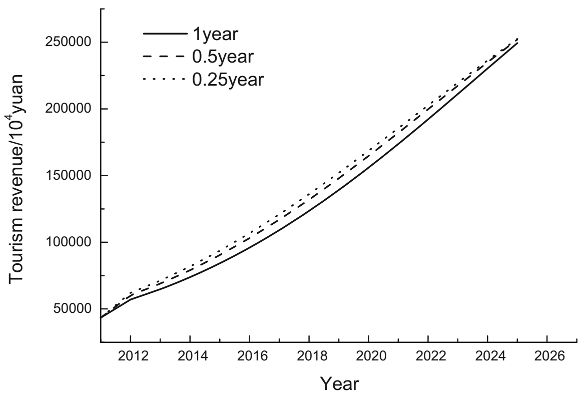


Fig. 4. Barkam's tourism revenues under different step simulation

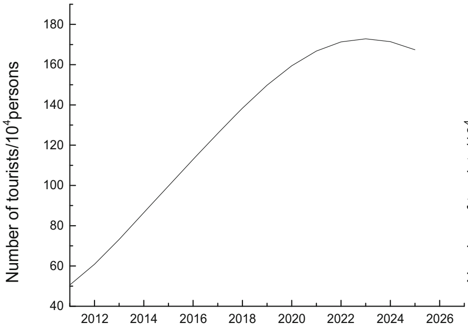


Fig. 5. Barkam's number of tourists simulation

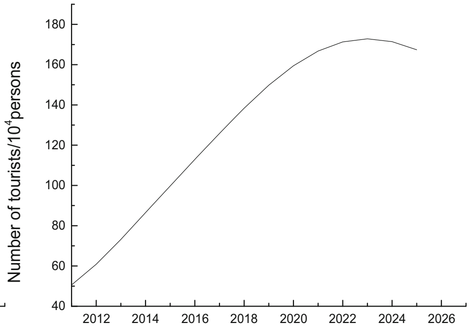


Fig. 6. Barkam's tourism revenue simulation

from 2011 to 2023, reaching 1.5 million by 2019 and having a peak at the end of 2023. After that, there will be a slight decline. The speculative reason is that pollution caused by the industrial clusters may affects the tourists' enthusiasm to travel to Barkam, while the development of the pan-tourism industry is driven by tourism. At the same time, Barkam's tourism revenue is also growing steadily and it will exceed 20 billion Yuan by 2022. We can see that the booming tourism industry of Barkam will be in a steady upward trend in the next decade from the simulation results. Therefore, in order to maintain the vitality of tourism industry and promote the development of industry clusters to propel urbanization, the government should pay more attention on the tourism industry meantime to concentrate on the impact of resource depletion and pollution from local industrial development [11].

(2) Simulation Analysis of Economy

As the Fig. 7 shown, the three industrial output values grow steadily in the period from 2011 to 2025. It is speculated that tertiary industry output values of Barkam will exceed 3 billion Yuan, while the GDP will exceed 4 billion Yuan.

Contrasting the primary and secondary industries, the tertiary industry output values increased significantly. The main reason is that tourism promote the development of related industries, such as transportation, lodging and catering industry and related social public services, so that the tertiary industry become the leading industry.

On the contrary, the growth rate of the primary industry is very slow, presumably due to the destruction of the local ecological environment and the obstruction of rural construction in the process of realizing urbanization. Local governments should take corresponding measures to strengthen the linkage of the tourism industry with the primary and secondary industry, promote the development of agricultural tourism and industrial tourism vigorously, extend the chain of business opportunities strenuously, and create an advantage platform to promote economic development effectively.

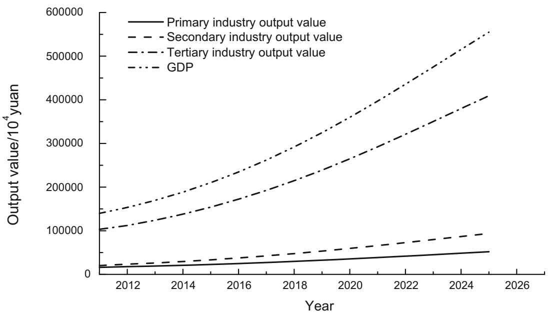


Fig. 7. Barkam's output value simulation

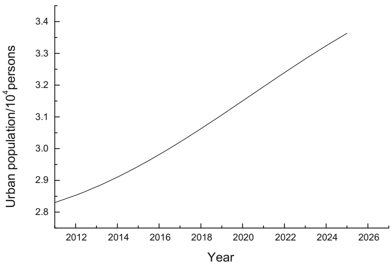


Fig. 8. Barkam's urban population simulation

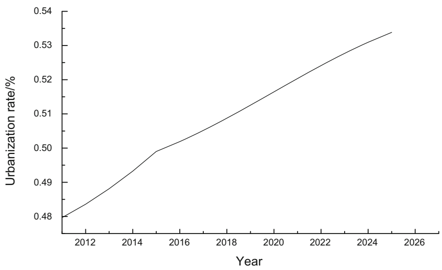


Fig. 9. Barkam's urbanization rate simulation

(3) Population Simulation

The simulation results of the employment population, the urban population and the urbanization rate are shown in Figs. 8 and 9. According to simulation results, the number of urban population will reach 33,000 while the urbanization rate of population will reach 53% by 2025, and both of them are in a steady rising state.

The main reason for the increase is the development of local tourism, driving the development of the tertiary industry such as services and circulation industries. These industries require low technical and easy to realize the labor force transfer.

Furthermore, the process of rural population to non-agricultural population is accelerated, and the in-situ urbanization of farmers is achieved which constantly improve the level of urbanization. From the data we can see that although the level of urbanization has increased year by year, the growth rate is relatively slow, mainly due to the fact that Barkam is located in the Tibetan area of northwest Sichuan, and the residents have long been engaged in agriculture and animal husbandry. The process of local urbanization has been speeded up by tourism, but some remote farmers still use the traditional agricultural mode, which affects the local urbanization process.

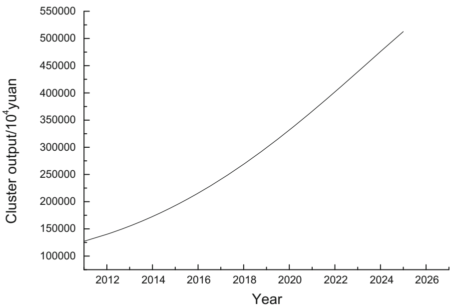


Fig. 10. Barkam's cluster output simulation

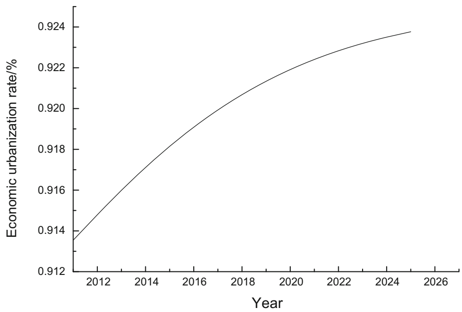


Fig. 11. Barkam's economic urbanization rate simulation

(4) Cluster Simulation

Cluster output value and economic urbanization rate simulation results shown in Figs. 10 and 11.

This paper refers to the definition of cluster output value by scholar Xiaoye Zhou [18], whose economic meaning is the development of secondary and tertiary industry in the region, relating to the entire country. As the Fig. 10 shows, the cluster output increased year by year from 2011 to 2025, which will exceed 5 billion Yuan in 2025. The stable increase in cluster output values means that the local secondary and tertiary industries are booming, which will lead to overall economic development of Barkam. Economic urbanization rate is defined as the proportion of cluster output values of GDP. The current urbanization rate of Barkam city has reached more than 90%, which shows that local industrial clusters have reached a high level. Figure 11 speculates that the Barkam's economic urbanization rate will be improved steadily in the next decade.

It can be concluded that the tourism industry of Barkam will develop rapidly in the next decade and the output values of industrial clusters will increase steadily. The urban GDP and urban population are also increasing. With the development of industrial clusters driven by tourism, the urbanization rate of Barkamis gradually increased, and the level of urbanization has laid a foundation for the local tourism development and industrial clusters. Meanwhile the development of tourism in Barkam has higher requirements for the whole urban system from the external to the internal connection service, which promotes the development of local construction, transportation, lodging catering and social public services, etc. The direct effect of tourism on the evolution of local urbanization has increased the attractiveness of cities and towns, and accelerated the process of Urbanization. In addition, we find that the growth rate of local agricultural output value is smaller than other two industries, which is not conducive to the coordinated development of various industries and the optimization of industrial structure. Therefore, the local government should maintain the smooth development of tourism and pay attention to the integration of tourism and related industries so that the development of the local three industries will be promoted harmoniously.

4 Conclusion

In this paper, the quantitative relationship was analyzed by SPSS, and the DYNAMO language and Vensim are used to establish the system dynamic model of pan-tourism urbanization. The three subsystems about tourism, industrial cluster and urbanization are interrelated through the visitor numbers, tourism revenue, and clusters output value and employment population. The interactive development relationship among the three reflects the industrial cluster driven by tourism development and the process of rural labor transfer. This model systematically analyzes pan-tourism urbanization operation mechanism, and it provides the corresponding scientific basis and theoretical support to promote urban and rural development, improve the level of characteristic urbanization and the construction of new towns. Because the selected city of Barkam has certain particularity, the system elements may not be comprehensive. In addition, as the difference of local development and economic level, the pan-tourism urbanization model may not be suitable for all cities and towns, so it is necessary to adjust specific elements and coefficients in a specific place in order to make a more scientific analysis.

Acknowledgements. This research was supported by the Key Funds of Sichuan Social Science Research Institution "System Science and Enterprise Development Research" (Grant NO. Xq15B09), the Youth Funds of Sichuan Provincial Education Department (Grant NO. 14ZB0014), and the Humanities Social and Sciences Research Funds of Education Ministry (Grant NO. 15XJC630001).

References

1. Alexandrova A, Vladimirov Y (2016) Tourism clusters in russia: what are their key features the case of vologda region. Worldwide Hospitality Tourism Themes 8(3):346–358
2. Eberlein RL, Peterson DW (1992) Understanding models with vensim. Eur J Oper Res 59:216–219
3. Fang C, Jing W (2013) A theoretical analysis of interactive coercing effects between urbanization and eco-environment. Chin Geogr Sci 23(2):147–162
4. Gimzauskiene E, Duoba K et al (2015) Tourism clusters in eastern poland - analysis of selected aspects of the operation. Procedia Soc Behav Sci 213:957–964
5. Jackson J (2006) Developing regional tourism in china: the potential for activating business clusters in a socialist market economy. Tourism Manag 27(4):695–706
6. Jackson J, Murphy P (2006) Clusters in regional tourism an australian case. Ann Tourism Res 33(4):1018–1035
7. Jian ML, Chi FL (2016) A qualitative study of urbanization effects on hotel development. J Hospitality Tourism Manag 29:135–142
8. Liu RX, Kuang J et al (2001) Principal component regression analysis with SPSS. Comput Methods Programs Biomed 71(2):141–147

9. Matinzadeh MM, Koupai JA et al (2017) Development of an innovative integrated model for the simulation of nitrogen dynamics in farmlands with drainage systems using the system dynamics approach. Ecol Model 347:11–28
10. Mullins P (1991) Tourism urbanization. Int J Urban Reg Res 15(3):326–342
11. Ning B, He Y (2007) Tourism development and water pollution: case study in Lijiang Ancient Town. China Population Resour Environ 17(5):123–127
12. Porter M (1990) Competitive Advantage of Nations. Free Press, New York
13. Qian J, Feng D, Zhu H (2012) Tourism-driven urbanization in china's small town development: a case study of zhapo town, 1986–2003. Habitat Int 36(1):152–160
14. Sun PJ, Song CL et al (2013) Non-coordination in china's urbanization: assessment and affecting factors. Chin Geogr Sci 23(6):729–739
15. Surya B (2016) The processes analysis of urbanization, spatial articulation, social change and social capital difference in the dynamics of new town development in the fringe area of makassar city (case study: In metro tanjungbunga area, makassar city). Procedia Soc Behav Sci 227:216–231
16. Tan Y, Xu H et al (2017) A study of best practices in promoting sustainable urbanization in china. J Environ Manag 193:8–18
17. Zhao B, Tang T, Ning B (2017) System dynamics approach for modelling the variation of organizational factors for risk control in automatic metro. Saf Sci 94:128–142
18. Zhou XY, Fu DM et al (2016) Research on interaction between industrial clusters and urbanization based on system dynamics. J Shenyang Univ Technol 9(1):47–52 (in Chinese)

A Study on the Three Different Channel Models in Supply Chain Under the Network Externalities

Jinjiang Yan, Jiankai Xing(✉), Kai Zhu, and Ke Liu

Business School, Sichuan University,
Chengdu 610065, People's Republic of China
mukaimu@163.com

Abstract. With the rapid developments of Internet, many manufacturers and retailers choose not only the offline channel but also the online channel. According to the choices of different channels, this paper models three types of channel models in supply chains. Considering the presence of the network externalities, the degree of substitution between online and offline channels and consumers' preference for two entities (the manufacture and the retailer), we study the effect of retailer's choice to online channel and the differences among the three different channel models. Our main findings are that compared to the first two models, the sales price of unit product will be smaller and in contrast the market size will be bigger in the third model. Also, when the entities make the best strategies, the retailer's profit and the total profit of the channel can reach the maximum. The study shows that the introduction of online channel can create more social values, which is why the dual-channel model of retailers is so widespread. In the end, a numerical example is given to verify the conclusions.

Keywords: Network externality · Supply chain · Online channel · Offline channel

1 Introduction

With the rapid development of the Internet, on the one hand, online sales have become a mode in which the manufacturers and retailers can increase sales. On the other hand, more and more products have gradually affected by the network externality. The introduction of online channel for the manufacturer can increase their sales and avoid the large retailers' monopoly of the offline channel. At the same time, it can provide opportunities for the retailer to expand the market and increase sales. Many well-known manufacturers (such as Lenovo, Apple, etc.) chose to introduce the online channel based on the offline channel. However, a lot of famous retailers (such as Alibaba, Wal-Mart Stores, etc.) had also introduced the online channel one after another. Nevertheless, most studies of dual-channels

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_36

focused on the choice of the manufacturer, the internal channel coordination and channel pricing, but ignored the impact of network externalities and the choice of retailers in dual-channel situation.

The network externality means that the customer can obtain more utilities when more suppliers are willing to join in the network and opposite situation is similar. In making strategies for channel markets, however, sellers (including the manufacturer and the retailer) have typically relied on assumptions and paradigms that apply to business without network effects. Therefore, the strategy that they make may not match the economics of their changing industries. Many literatures studied the pricing strategy for the dual-channel. Liu [13] pointed that compared to the traditional single channel, when the manufacturer brings in the online channel, pricing strategies will have different effects on the channel. Leider [9] studied experimentally bargaining in a multiple-tier supply chain with horizontal competition and sequential bargaining between tiers. They found that the structural issue of cost differentials dominates personal characteristics in explaining outcomes, with profits in a tier generally increasing with decreased competition in the tier and increasing with decreased competition in alternate tiers. Naeem, Charles and Rafay [3] considered a dynamic problem of joint pricing and production decisions for a profit-maximizing firm that produced multiple products. They presented a solution approach which is more general than previous approaches that require the assumption of a specific demand function. Chuang and Yang [7] modeled the deteriorating goods pricing strategy with reference price effects and solved the dynamic problem by obtained theoretical results.

Some scholars carried out their research from the aspect of dual-channel coordination. Shamir [14] studied the operational motivation of a retailer and showed that by making forecast information publicly available to both his manufacturer and to the competitor, a retailer is able to credibly share his forecast information-an outcome that cannot be achieved by merely exchanging information within the supply chain. Yang and Zhuo etc. studied a two-period supply chain that consists of a retailer and a supplier and found that a revenue sharing contract significantly affects the retailer's payment behavior and supplier's wholesale price [16]. Zhong, Xi and Jing [17] analyzed the coordinating mechanisms for a single-period supply chain comprising of one supplier and one retailer. They pointed that the transfer payment plus a contract could mitigate the down side risk effect on the supply chain performance. Yue and Allen, etc. [15] studied a supply chain consisting of one supplier and n retailers and pointed the pricing strategy under different situations. Wenyi et al. [6] presented an integrated model of marketing choices and supply chain design decisions and developed the closed-loop supply chain. Yulan and Stein, etc. reviewed operational models in service supply chain management (SSCM) and examined the definitions of SSCM [6]. Panda et al. [12] analyzed coordination of a manufacturer-distributer-retailer supply chain, where the manufacturer exhibited corporate social responsibility [12]. They built a manufacturer-Stackelberg game setting to propose a contract-bargaining process and resolved channel conflict. Jing, Hui and Ying [5] analyzed the channel effects of four kinds of dual-channels' pricing strategies. Aimin and Liwen analyzed the stochastic demand, joint promotion, the price competition and coordination problems between the manufacturer and the retailer [1]. Ravi and Rajib [11] studied the co-opetition between differentiated platform in two-sided markets. In their study, they highlighted the importance of technology. They pointed that collaboration might provide incentives for a dominant platform. In this paper, we consider the platform/channel ownership is independent, not belonging to the seller or the consumer.

Some scholars focus on the study of network externalities for the competitive strategy. Blocher [4] showed that amplifying network externalities among mutual funds could explain substantial flow-based effects documented in the literature. Liu [10] identified the valid mechanism for the alternative range of profit-sharing contracts and analyzed the effect of product substitutability coefficient and network externalities on the alliance and profit-sharing contract. Hagiu and Spulber [8] presented first-party content and coordination in two-sided markets and they found that the strategic use of first-party content by two-sided platform was driven by two key factors: the nature of buyer and seller expectations. Edward and Anderson researched the platform performance investment in the presence of network externalities [2]. They carried out a full analysis of three distinct settings: monopoly, price-setting duopoly and pricing-taking duopoly. They concluded that the conditions under which offering a platform with lower performance but greater availability of content can be a winning strategy.

These literatures have analyzed the pricing and coordination strategies of the dual-channels and the influence of network externalities on pricing and competitive strategy of enterprises respectively. We can conclude that the main problem of pricing and coordination is to coordinate the distribution of benefits between manufacturers and retailers in the structure of the manufacturer's choice of dual channel. The network externality can increase the strength of the sharing economy, bring about a positive cycle effects, and provide guidance for pricing and coordination of enterprises. However, from the two aspects of researches, there exists the following deficiencies.

The retailer's behavior can also have a great impact on the dual-channel in the supply chain system. But most of the studies about the dual-channel and the network externalities are limited to the behavior of the manufacturer. In addition, the network externalities also have a great influence on the pricing and decision making. Many literatures about the dual-channel have largely ignored the effect of network externalities on consumers and the whole supply chain system. Many studies about the network externalities simply consider the effect of network externalities on the single channel mode, not using it to dual-channel mode.

With the premise that both entities' (including the manufacturer and the retailer) optimal profits must greater than zero, this paper considers the products with the network externalities and builds three channel models under the condition that the retailer is the leader in the structure models. This paper names them as single channel model (SCM), the dual-channel model of manufacturers

(DCMM) and the dual-channel model of retailers (DCMR) respectively. First, this paper analyzes the optimal product pricing and strategy for the manufacturer and the retailer under three different channel models. Then we analyze the differences among the three models and their effect on the local supply chain. At last, this paper uses the numerical examples to demonstrate the conclusions.

2 Model Assumptions and Notations Description

Assumed that the manufacturer only produces a kind of product with network externality. Considering that the sale period is not too long, consumers only buy the goods from the manufacturer or the retailer on the channel and there is no intersection between the consumption of different channels.

The retailer is the leader in the structure in all models. In the single channel model, firstly, according to the retailer's own marginal profit m_1, the manufacturer decides the wholesale price ω_1 to maximize their own profits. Then, the retailer adjusts its revenue to maximize its own marginal profit according to the wholesale price. Further, the retailer sets the sales price of the product $p_1 = \omega_1 + m_1$. In the dual-channel model of manufacturers, firstly, according to the retailer's own marginal profit m_2, the manufacturer decides the wholesale price ω_1 and the sale price p_{22} on online channel to maximize their own profits. Second, the retailer adjusts its revenue to maximize its own margin according to the wholesale price. At last, the retailer sets the sales price of the product. Similar to the dual-channel model of manufacturers, in the dual-channel model of retailers, the manufacturer decides the wholesale price and the sale price ω_2 on online channel p_{22}. The main difference is that the retailer needs to decide the sale price on online channel p_{22}.

Consumers can buy goods through different channels, but they will have different utilities and the estimated value of the product on offline or online channels will also be different. For example, In the physical store, consumers can immediately get product information and then purchase products. While through the network channel, consumers can get more products no matter how far it is. With the development of modern logistics, waiting time will become very short. So we denote the degree of substitution or monopoly between online and offline channels by $\beta\,(0 \le \beta < 1)$, so $1-\beta$ indicates the degree that the online channel cannot monopolize. While refers to the manufacturer's credit tendency in the mind of the consumer, so does the $1-\gamma$.

In the single channel model, the consumer's estimation value of the product on the offline channel is V_{11}; in the dual-channel model of manufacturers, the consumer's estimation value of the product on the offline and online channels are V_{21}, V_{22}; in the dual-channel model of retailers, the consumer's estimation value of the product on the offline channel is V_{31}, while the consumer's estimation value of the manufacturer's and the retailer's product on the online channel are V_{32}, V_{33}.

The consumer will select the channel only when their utilities are greater than zero. Without loss of generality, assume the customer is heterogeneous and subject to uniform distribution. In model one, θ_{11} refers to the consumers' preference distribution on offline channel. It represents the consumer's measure for each potential transaction; In model two, we use θ_{21} instead of θ_{11} and denote the consumer's preference distribution for manufacturers of online channel by θ_{22}. In model three, we use θ_{31}, θ_{32} instead of θ_{21}, θ_{22} and denote the consumer's preference distribution for manufacturers of online channel by θ_{32}.

The sales cost of unit product for the manufacturer on offline channel can be interpreted by c_1, while for the retailer is c_2. Because the cost of online channel is lower than the offline channel, so define k_1 as the manufacturer's allocation efficiency relative to the offline channel, while for the retailer is k_2. The smaller the k, the lower the cost. To be more representative, we can get $k \in (0, 1)$. α_1 expresses the strength of network effects of consumers who select retailer's offline channel; α_2 expresses the strength of network effects of consumers who select manufacturer's online channel; α_3 expresses the strength of network effects of consumers who select retailer's online channel. Without loss of generality, $\alpha \in (0, 1)$.

In the single channel model, A_{11} represents the largest potential market demand on offline channel, N_{r1}denotes the number of retailers on offline channel. In the dual-channel model of manufacturers, use A_{21}, N_{r2} instead of A_{11}, N_{r1}, A_{22} represents the largest potential market demand on online channel, N_{s2} denotes the number of manufacturers on online channel. Due to the widening of the market, we can conclude $A_{21} + A_{22} > A_{11}$. At the same time, because the manufacturer adds the online channel, and appeal some consumers, so it will inevitably lead to a decrease in the number of retailers. So $N_{r2} \le N_{r1}$. In the dual-channel model of manufacturers, similar to model two, A_{31} can instead of A_{21}; A_{32} represents the largest potential market demand for the manufacturer on online channel; A_{33} represents the largest potential market demand for the retailer on online channel; denotes the number of retailers on online channel; N_{s3} denotes the number of manufacturers on online channel; N_{r4} denotes the number of retailers on offline channel; Similarly, $A_{32} + A_{33} > A_{22}$, $N_{s3} \le N_{s2}$ and $N_{r4} \le N_{r2}$.

3 The Model

3.1 The Single Channel Model (SCM)

Based on the above assumptions, we study the single channel model firstly. The supply chain mode is shown in Fig. 1.

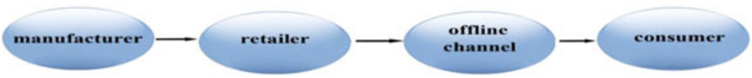


Fig. 1. The single channel mode (SCM)

Assume that the utility function of the consumer on the single channel is:

$$U_{11} = V_{11} + N_{r1}\theta_{11}\alpha_1 - p_{11}. \tag{1}$$

According to Eq. (1), for this channel, the consumer's individual rationality (participation) constraint is:

$$U_{11} \geq 0. \tag{2}$$

So

$$\theta_{11} \geq \frac{p_{11} - V_{11}}{N_{r1}\alpha_1}. \tag{3}$$

As a result, according to Eq. (3), the number of consumers who join the retailer's offline channel is:

$$D_{11} = A_{11}\int_{\frac{p_{11}-V_{11}}{N_{r1}\alpha_1}}^{1} d\theta = \frac{(N_{r1}\alpha_1 - p_{11} + V_{11})A_{11}}{N_{r1}\alpha_1}. \tag{4}$$

And it must satisfy:

$$D_{11} \geq 0. \tag{5}$$

(1) The manufacturer's profit is:

$$\varPi_{M1} = (\omega_1 - c_1)D_{11} = A_{11}(\omega_1 - c_1)\frac{N_{r1}\alpha_1 - \omega_1 - m_1 + V_{11}}{N_{r1}\alpha_1}. \tag{6}$$

Proposition 1. *The manufacturer's profit function is a concave function of the wholesale price, so when the manufacturer's profit reaches the maximum, the price's optimal value is:*

$$\omega_1^* = \frac{N_{r1}\alpha_1 - m_1 + c_1 + V_{11}}{2}. \tag{7}$$

(2) The retailer's profit of offline platform is:

$$\varPi_{\mathrm{R1}} = (p_{11} - \omega_1 - c_2)D_{11} = A_{11}(m_1 - c_2)\frac{N_{r1}\alpha_1 - m_1 - c_1 + V_{11}}{2V_{11}\alpha_1}. \tag{8}$$

Proposition 2. *The retailer's profit function is a concave function of the earning of unit product on offline channel, so when the retailer's profit reaches the maximum, the margin's optimal value is:*

$$m_1^* = \frac{N_{r1}\alpha_1 - c_1 + c_2 + V_{11}}{2}. \tag{9}$$

In the single channel model, the final optimal values for these parameters are as following:

$$\omega_1^* = \frac{N_{r1}\alpha_1 + 3c_1 - c_2 + V_{11}}{4}, \tag{10}$$

$$p_{11}^* = \frac{3N_{r1}\alpha_1 + c_1 + c_2 + 3V_{11}}{4}, \tag{11}$$

$$\Pi_{M1}^* = \frac{(N_{r1}\alpha_1 - c_1 - c_2 + V_{11})^2 A_{11}}{16N_{r1}\alpha_1}, \tag{12}$$

$$\Pi_{R1}^* = \frac{(N_{r1}\alpha_1 - c_1 - c_2 + V_{11})^2 A_{11}}{8N_{r1}\alpha_1}, \tag{13}$$

$$\Pi_1^* = \Pi_{M1}^* + \Pi_{R1}^* = \frac{3(N_{r1}\alpha_1 - c_1 - c_2 + V_{11})^2 A_{11}}{16N_{r1}\alpha_1}. \tag{14}$$

3.2 The Dual-Channel Model of Manufacturers (DCMM)

Based on the above assumptions and the first model, we study the dual-channel model of manufacturers next. The supply chain mode is shown in Fig. 2.

Fig. 2. The dual-channel model of manufacturers (DCMM)

As consumers have their own preferences for online and offline channels, assume that the utility function of the consumer on the offline channel and online channel respectively are:

$$U_{21} = V_{21} + N_{r2}\theta_{21}\alpha_1 - p_{21} - (1-\beta)V_{21}V_{22}, \tag{15}$$

$$U_{22} = V_{22} + N_{s2}\theta_{22}\alpha_2 - p_{22} - \beta V_{21}V_{22}. \tag{16}$$

On the basis of Eqs. (15) and (16), the consumer's individual rationality (participation) constraints for these two channels are:

$$U_{21} \geq 0, \tag{17}$$

$$U_{22} \geq 0. \tag{18}$$

So

$$\theta_{21} \geq \frac{p_{21} + (1-\beta)V_{21}V_{22} - V_{21}}{N_{r2}\alpha_1}, \tag{19}$$

$$\theta_{22} \geq \frac{p_{22} + \beta V_{21}V_{22} - V_{22}}{N_{s2}\alpha_2}. \tag{20}$$

As a result, according to Eqs. (19) and (20), the numbers of consumers who join retailer's offline channel and online channel respectively are:

$$D_{21} = A_{21}\frac{N_{r2}\alpha_1 - p_{21} - (1-\beta)V_{21}V_{22} + V_{21}}{N_{r2}\alpha_1}, \tag{21}$$

$$D_{22} = A_{22}\frac{N_{s2}\alpha_2 - p_{22} - \beta V_{21}V_{22} + V_{22}}{N_{s2}\alpha_2}. \tag{22}$$

And they must satisfy:

$$D_{21} \geq 0, \tag{23}$$

$$D_{22} \geq 0. \tag{24}$$

The manufacturer decides the wholesale price and the online channel's price of unit product to maximize their own profits. The manufacturer's profit is:

$$\Pi_{M2} = (\omega_2 - c_1)D_{21} + (p_{22} - k_1c_1)D_{22}. \tag{25}$$

Proposition 3. *The manufacturer's profits function is a concave function of the wholesale price and the online channel's price of unit product, so when the manufacturer's profit reaches the maximum, the optimal values of the prices are:*

$$p_{22}^* = \frac{N_{s2}\alpha_2 - \beta V_{21}V_{22} + k_1c_1 + V_{22}}{2}, \tag{26}$$

$$\omega_2^* = \frac{N_{r2}\alpha_1 - (1-\beta)V_{21}V_{22} - m_2 + c_1 + V_{21}}{2}. \tag{27}$$

At the same time, the retailer's decision-making is the same as the single channel model. The retailer's profit of offline platform is:

$$\begin{aligned}\Pi_{R2} &= (p_{21} - \omega_2 - c_2)D_{21}\\ &= A_{21}\frac{m_2 - c_2}{2}\frac{N_{r2}\alpha_1 - (1-\beta)V_{21}V_{22} - m_2 - c_1 + V_{21}}{N_{r2}\alpha_1}.\end{aligned} \tag{28}$$

Proposition 4. *The retailer's profits function is a concave function of the earning of unit product on offline channel, so when the retailer's profit reaches the maximum, the margin's optimal value is:*

$$m_2^* = \frac{N_{r2}\alpha_1 - (1-\beta)V_{21}V_{22} - c_1 + c_2 + V_{21}}{2}. \tag{29}$$

So in the dual-channel model of manufacturers, the final optimal values for these parameters are as the following:

$$\omega_2^* = \frac{N_{r2}\alpha_1 - (1-\beta)V_{21}V_{22} + 3c_1 - c_2 + V_{21}}{4}, \tag{30}$$

$$p_{21}^* = \frac{3N_{r2}\alpha_1 - 3(1-\beta)V_{21}V_{22} + c_1 + c_2 + 3V_{21}}{4}, \tag{31}$$

$$p_{22}^{*} = \frac{N_{s2}\alpha_2 - \beta V_{21}V_{22} + k_1 c_1 + V_{22}}{2}, \tag{32}$$

$$\begin{aligned} \Pi_{M2}^{*} = & A_{21}\frac{\left[N_{r2}\alpha_1 - (1-\beta)V_{21}V_{22} - c_1 - c_2 + V_{21}\right]^2}{16N_{r2}\alpha_1} \\ & + A_{22}\frac{\left(N_{s2}\alpha_2 - \beta V_{21}V_{22} - k_1 c_1 + V_{22}\right)^2}{4N_{s2}\alpha_2}, \end{aligned} \tag{33}$$

$$\Pi_{R2}^{*} = A_{21}\frac{\left[N_{r2}\alpha_1 - (1-\beta)V_{21}V_{22} - c_1 - c_2 + V_{21}\right]^2}{8N_{r2}\alpha_1}. \tag{34}$$

Finally, consider the prerequisite: the manufacturer's profit in this model should be greater than in model one (Because only in this condition, the manufacturer will choose to introduce online channel). So for the manufacturer, their optimal profit must be satisfied:

$$\Pi_{M2} \geq \Pi_{M1}. \tag{35}$$

Let $\frac{(N_{r1}\alpha_1 - c_1 - c_2 + V_{11})^2 A_{11}}{16N_{r1}\alpha_1} = C_1$, $A_{21}\frac{[N_{r2}\alpha_1 - (1-\beta)V_{21}V_{22} - c_1 - c_2 + V_{21}]^2}{16N_{r2}\alpha_1} = C_2$ and $\beta V_{21}V_{22} - k_1 c_1 + V_{22} = C_3$. So according to formula (35), we can get:

$$\alpha_2 \geq \frac{\frac{4(C_1 - C_2)}{A_{22}} + 2N_{s2}C_3 + \sqrt{\left[\frac{4(C_1 - C_2)}{A_{22}} + 2N_{s2}C_3\right]^2 - 4N_{s2}^2 C_3}}{2N_{s2}^2}. \tag{36}$$

3.3 The Dual-Channel Model of Retailers (DCMR)

In the end, we analyze the dual-channel model of retailers. According to the above analysis, the supply chain mode is shown in Fig. 3.

Fig. 3. The dual-channel model of retailers (DCMR)

In the case of this model, consumers have not only different choices of channels, but also preferences for the entities in the channels. Therefore, the net utilities to a consumer on the retailer's offline platform, the manufacturer's online platform, the retailer's online platform are given by:

$$U_{31} = V_{31} + N_{r4}\theta_{31}\alpha_1 - p_{31} - (1-\beta)V_{31}(V_{32} + V_{33}), \tag{37}$$

$$U_{32} = V_{32} + N_{s3}\theta_{32}\alpha_2 - p_{32} - \beta\gamma V_{31}(V_{32} + V_{33}). \tag{38}$$

Similarly, the numbers of consumers who join retailer's offline channel, manufacturer's online channel and retailer's online channel respectively are:

$$D_{31} = A_{31}\frac{N_{r4}\alpha_1 - p_{31} - (1-\beta)V_{31}(V_{32}+V_{33}) + V_{31}}{N_{r4}\alpha_1}, \tag{39}$$

$$D_{32} = A_{32}\frac{N_{s3}\alpha_2 - p_{32} - \beta\gamma V_{31}(V_{32}+V_{33}) + V_{32}}{N_{s3}\alpha_2}, \tag{40}$$

$$D_{33} = A_{33}\frac{N_{r3}\alpha_3 - p_{33} - \beta(1-\gamma)V_{31}(V_{32}+V_{33}) + V_{33}}{N_{r3}\alpha_3}. \tag{41}$$

According to the actual situation, they must satisfy:

$$D_{31} \geq 0, \tag{42}$$

$$D_{32} \geq 0, \tag{43}$$

$$D_{33} \geq 0. \tag{44}$$

Similar to the model two, the manufacturer decides the wholesale price and the online channel's price of unit product to maximize their own profits. The manufacturer's profit in this model is:

$$\Pi_{M3} = (\omega_3 - c_1)D_{31} + (p_{32} - k_1 c_1)D_{32}. \tag{45}$$

Proposition 5. *The manufacturer's profits function is a concave function of the wholesale price and the online channel's price of unit product, so when the manufacturer's profit reaches the maximum, the optimal values of these prices are:*

$$p_{32}^* = \frac{N_{s3}\alpha_2 - \beta\gamma V_{31}(V_{32}+V_{33}) + k_1 c_1 + V_{32}}{2}, \tag{46}$$

$$\omega_3^* = \frac{N_{r4}\alpha_1 - (1-\beta)V_{31}(V_{32}+V_{33}) - m_3 + c_1 + V_{31}}{2}. \tag{47}$$

Different from the model two, the retailer not only decides the earning of unit product on offline channel, but also the online channel's price of unit product. The retailer's total profit in this model is:

$$\Pi_{R3} = (p_{31} - \omega_3 - c_2)D_{31} + (p_{33} - k_2 c_2)D_{33}. \tag{48}$$

Proposition 6. *The retailer's profits function is a concave function of the earning of unit product on offline channel and the online channel's price of unit product, so when the retailer's profit reaches the maximum, the optimal values are:*

$$m_3^* = \frac{N_{r4}\alpha_1 - (1-\beta)V_{31}(V_{32}+V_{33}) + c_2 - c_1 + V_{31}}{2}, \tag{49}$$

$$p_{33}^* = \frac{N_{r3}\alpha_3 - \beta(1-\gamma)V_{31}(V_{32}+V_{33}) + k_2 c_2 + V_{33}}{2}. \tag{50}$$

Therefore, in the dual-channel model of retailers, the final optimal values for these parameters are as the following:

$$
\begin{aligned}
\omega_3^* &= \frac{N_{r4}\alpha_1 - (1-\beta)V_{31}(V_{32}+V_{33}) - c_2 + 3c_1 + V_{31}}{4}, && (51)\\
p_{31}^* &= \frac{3N_{r4}\alpha_1 - 3(1-\beta)V_{31}(V_{32}+V_{33}) + c_2 + c_1 + 3V_{31}}{4}, && (52)\\
p_{32}^* &= \frac{N_{s3}\alpha_2 - \beta\gamma V_{31}(V_{32}+V_{33}) + k_1c_1 + V_{32}}{2}, && (53)\\
p_{33}^* &= \frac{N_{r3}\alpha_3 - \beta(1-\gamma)V_{31}(V_{32}+V_{33}) + k_2c_2 + V_{33}}{2}, && (54)\\
\Pi_{M3}^* &= A_{33}\frac{[N_{r4}\alpha_1 - (1-\beta)V_{31}(V_{32}+V_{33}) - c_1 - c_2 + V_{31}]^2}{16N_{r4}\alpha_1}\\
&\quad + A_{34}\frac{[N_{s3}\alpha_2 - \beta\gamma V_{31}(V_{32}+V_{33}) - k_1c_1 + V_{32}]^2}{4N_{s3}\alpha_2}, && (55)\\
\Pi_{R3}^* &= A_{33}\frac{[N_{r4}\alpha_1 - (1-\beta)V_{31}(V_{32}+V_{33}) - c_1 - c_2 + V_{31}]^2}{8N_{r4}\alpha_1}\\
&\quad + A_{35}\frac{[N_{r3}\alpha_3 - \beta(1-\gamma)V_{31}(V_{32}+V_{33}) - k_2c_2 + V_{33}]^2}{4N_{r3}\alpha_3}. && (56)
\end{aligned}
$$

Finally, consider the prerequisite: the manufacturer's profit in this model should be greater than in model one and the retailer's profit in this model should be greater than in model two. So for the manufacturer, their optimal profit must be satisfied:

$$\Pi_{M3} \geq \Pi_{M1}. \tag{57}$$

Let $A_{33}\frac{[N_{r4}\alpha_1-(1-\beta)V_{33}(V_{34}+V_{35})-c_1-c_2+V_{33}]^2}{16N_{r4}\alpha_1} = C_4$ and $\beta\gamma V_{31}(V_{32}+V_{33}) + k_1c_1 - V_{32} = C_5$. So:

$$\alpha_2 \geq \frac{4(\frac{C_1-C_4}{A_{32}}) + 2N_{s3}C_5 + \sqrt{[4(\frac{C_1-C_4}{A_{32}}) + 2N_{s3}C_5]^2 - 4N_{s3}^2C_5}}{2N_{s3}^2}. \tag{58}$$

At last, for the retailer, their optimal profit must be satisfied:
Let $\beta(1-\gamma)V_{31}(V_{32}+V_{33}) + k_2c_2 - V_{33} = C_6$. According to formula (57)

$$\alpha_3 \geq \frac{8(\frac{C_2-C_4}{A_{33}}) + 2N_{r3}C_6 + \sqrt{[8(\frac{C_2-C_4}{A_{33}}) + 2N_{r3}C_6]^2 - 4N_{r3}^2C_6}}{2N_{r3}^2}. \tag{59}$$

Comparing the three models and when these prerequisites are met, we can get the following conclusions. With manufacturers and retailers choose the online channel in turn, it can be seen that the price of unit product on offline channel gradually reduces. In model three, the online channel's price of unit product for the manufacturer and the retailer both lower than the online channel's price in model two. With the introduction of online channel, the wholesale price is gradually reduced and the number of consumers who join retailer's offline channel will decrease. As the retailer chooses the online channel, the number of consumers who join manufacturer's online channel will also decrease, but the total number of consumers will increase. Compared to the model one, when the manufacturer chooses the online channel, their profit gets to the maximum, then when the retailer uses online channel, the manufacturer's profit reduced, but is higher than in the model one. In model two, when the manufacturer chooses the online channel, the retailer's profit gets to the minimum, then when they choose the online channel, their profits will increase and higher than in model one. Along with the introduction of the online channel, the overall profit of the channel will increase.

4 Numerical Example

This section presents numerical simulations to clarify the findings which is not easy to understand. First, we maintain the following assumptions over the simulations. For the case of the single channel model, in order to achieve profitability, some variables need to meet some restrictions: $A_{11} \geq p_{11} \geq \omega_1 \geq c_1$. In model two, in order to prevent retailers ordering goods from online channels and be able to make profit, they need to meet $A_{21} \geq p_{21} \geq \omega_2 \geq c_1$, $p_{22} \geq \omega_2 \geq c_1$ and $A_{22} \geq p_{22} \geq k_1 c_1$. Similar to model two, in model three, $A_{31} \geq p_{31} \geq \omega_3 \geq c_1$, $p_{32} \geq \omega_3 \geq c_1$, $A_{32} \geq p_{32} \geq k_1 c_1$, $p_{33} \geq \omega_3 \geq c_1$ and $A_{33} \geq p_{33} \geq k_2 c_2$.

According to formula $A_{11} \geq p_{11} \geq \omega_1 \geq c_1$, $A_{21} \geq p_{21} \geq \omega_2 \geq c_1$, $A_{31} \geq p_{31} \geq \omega_3 \geq c_1$ and formulas (2), (21), (22), (42), (43) and (44), the prerequisite is:

$$\alpha_1 \geq \frac{c_1 + c_2 + (1-\beta)V_{31}(V_{32}+V_{33}) - V_{31}}{N_{r4}}. \tag{60}$$

According to formula $p_{22} \geq \omega_2 \geq c_1$, $A_{22} \geq p_{22} \geq k_1 c_1$, $p_{32} \geq \omega_3 \geq c_1$, $A_{32} \geq p_{32} \geq k_1 c_1$ and formulas (21), (22), (36), (42), (43), (44) and (57), the prerequisite is:

$$\alpha_2 \geq \frac{N_{r4}\alpha_1 + (\beta - 1 + 2\beta\gamma)V_{31}(V_{32}+V_{33}) + (3-2k_1)c_1 - c_2 + V_{31} - 2V_{32}}{2N_{s3}}. \tag{61}$$

According to formula $p_{32} \geq \omega_3 \geq c_1$, $A_{32} \geq p_{32} \geq c_1$, $p_{32} \geq p_{32} \geq k_1 c_1$ and formulas (42), (43) and (44), and controlling other variables constant, the prerequisite is:

$$\beta \leq \frac{2N_{r3}\alpha_3 - N_{r4}\alpha_1 + V_{31}(V_{32}+V_{33}) + (2k_2+1)c_2 - 3c_1 + 2V_{33} - V_{31}}{(3-2\gamma)V_{31}(V_{32}+V_{33})}. \tag{62}$$

So according to previous assumptions, these parameters' values can be set: $A_{11} = 16$, $V_{11} = 0.35$, $\alpha_1 = 0.7$, $c_1 = 0.5$, $c_2 = 0.3$, $N_{r1} = 10$, $A_{21} = 14$, $A_{22} = 13$, $V_{21} = 0.5$, $V_{22} = 0.37$, $\alpha_2 = 0.6$, $\beta = 0.7$, $k_2 = 0.5$, $N_{r2} =$, $N_{s2} = 0.6$, $A_{32} = 10$, $A_{31} = 13$, $A_{32} = 12.5$, $V_{31} = 0.6$, $V_{31} = 0.6$, $V_{32} = 0.4$, $V_{33} = 0.5$, $\gamma = 0.6$,

$k_2 = 0.1$, $\alpha_3 = 0.5$, $N_{s3} = 5$, $N_{s3} = 5.5$ and $N_{r4} = 7$. According to formulas (60), (61) and (62), and controlling other variables constant, these prerequisites are $0.052 \le \alpha_1 \le 1$, $0.569 \le \alpha_2 \le 1$, and $0 \le \beta \le 0.782$. So the prices of unit product on different channels can be shown in Fig. 4.

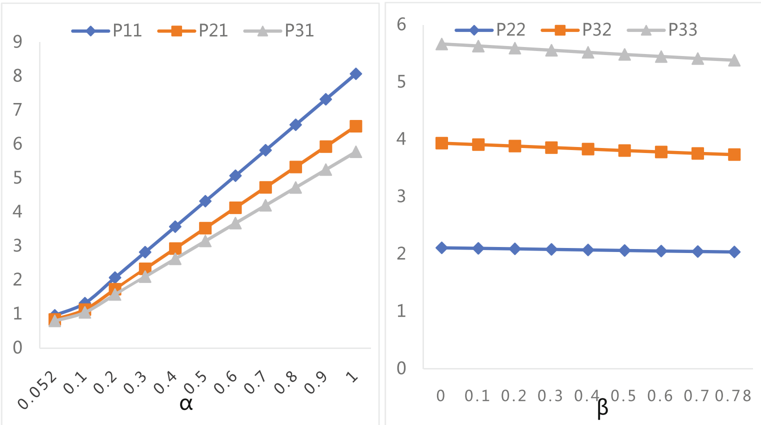


Fig. 4. The prices of unit product

As shown in Fig. 4, under the premise that manufacturers and retailers can choose the online channel, we can conclude that the price of the retailer's unit product on offline channel gradually decrease; the manufacturer's and the retailer's online channel's prices of unit product are also lower than in single channel model. That means consumers are very favorable when the manufacturer and the retailer both choose the online channel.

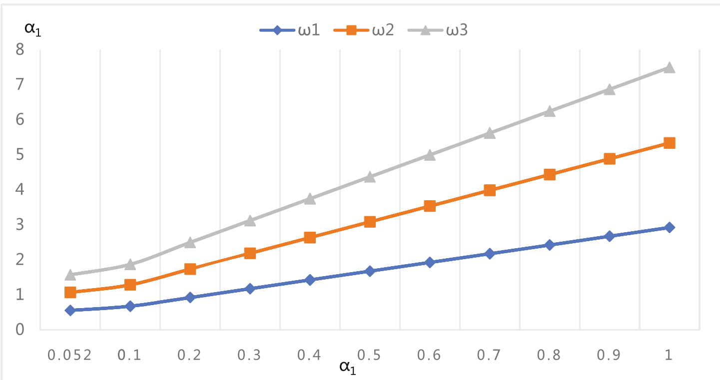


Fig. 5. The wholesale price of unit product

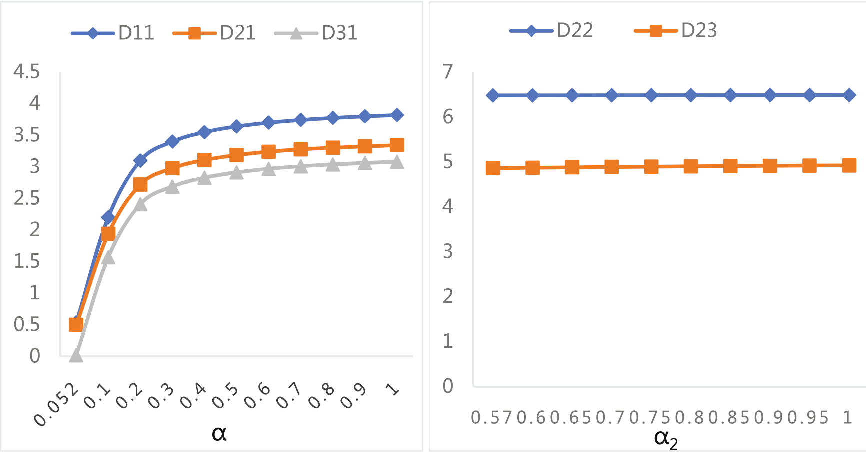


Fig. 6. The number of consumers

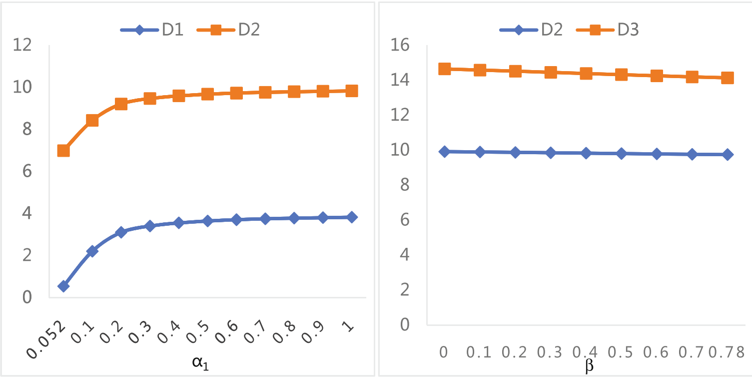


Fig. 7. The number of consumers

Compared the three models, the wholesale price of unit product is gradually reduced with the introduction of online channel. It's benefit for the retailer (Figs. 6 and 7).

With the increase of the network externality, the number of consumers who join the offline and online channels are all increasing in the three modes. Compared the three modes, due to the introduction of online channel, some consumers transfer to online channel. So the number of consumers who join the offline channel will decrease. When the retailer chooses to introduce the online channel, the number of consumers who join the manufacturer's online channel will also decrease. As shown in Fig. 7, the total number of consumers is increasing with the increase of network externalities and decreasing with the degree of substitution or monopoly between online and offline channels. With the introduction of online channel, the total number of consumers is increasing in turn. The

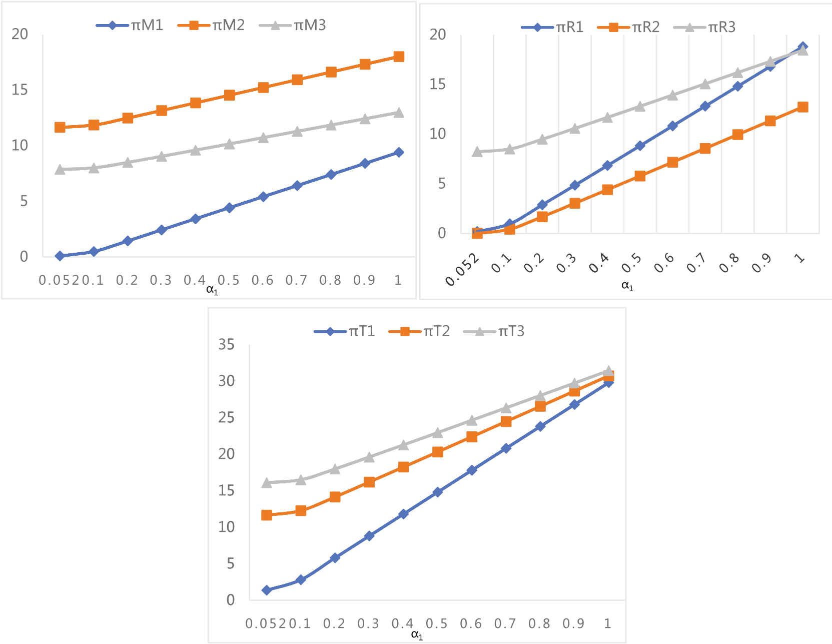


Fig. 8. The profit

total number of consumers can increase with the introduction of online channel. At the same time, consumers have more channels to choose, manufacturers and retailers have more demand for products, and thus these behaviors can increase consumption and create more social values.

We can see from Fig. 8, with the increase of the network externalities all profits are increasing. In model one, the profit of the manufacturer is the lowest. In model two, as the manufacturer chooses the online channel their profits reach the maximum. In model three, as the retailer chooses the online channel, their profits somewhat reduce but higher than in model one. On the contrast, in model one, the profit of the retailer is lower than in model three but higher than model one. In model two, as the manufacturer chooses the online channel, their profits reach the minimum. In model three, as the retailer chooses the online channel, their profits increase and reach the maximum. The total profit is increasing with the introduction of online channel. That means when manufacturers and retailers both choose to use online channel, they will create more social values.

5 Conclusions

This paper considers the choice of the retailer based on the choice of the manufacturer under the network externalities. We focus on the influence when the entities have different choices. By contrasting the single channel model, the dual-channel model of manufacturers and the dual-channel model of retailers, the introduction of online channel is benefit to the local supply chain.

Our study of channel performance under the affecting of the network externalities has yielded several important insights that should be of managerial interest. There are certain conditions for the manufacturer's and retailer's selection of the online channel. When manufacturers or retailers introduced the online channel first, the other one will make the same choice. With the introduction of online channel, the whole price will decrease but the number of consumers and the total profit will increase. That is why the dual-channel model of retailers is so widespread.

In developing our models, there are a few simplifying assumptions. For example, consistent with many literatures, we only study a single-period model and assume the consumer can only choose one channel. While we believe the intuition of our results applies to cases with different assumption, future research might investigate how such assumptions (e.g., independent of retailers' utility) affect the conditions in the proposition. Finally, future research could divide the channel into monopoly channel and competitive channel, so we can further discuss the choice of the manufacturer and the retailer between two channels or among three channels. There also might conduct a comprehensive analysis of manufacturers' or retailers' decisions which include different first-party content, technology and market share.

Acknowledgements. This research is funded by The Soft Science Research Project of Sichuan Province (No. 2016ZR0042).

References

1. Ai-Min YU, Liu LW (2012) Competing and coordination strategies for the dual channel under stochastic demand and cooperative promotion. J Ind Eng Eng Manag 26(1):151–155
2. Anderson EG, Parker GG, Tan B (2014) Platform performance investment in the presence of network externalities. Inf Syst Res 25(1):152–172
3. Bajwa N, Sox CR et al (2016) Coordinating pricing and production decisions for multiple products. Omega 64:86–101
4. Blocher J (2016) Network externalities in mutual funds. J Finan Mark 30:1–26
5. Chen J, Zhang H, Sun Y (2012) Implementing coordination contracts in a manufacturer stackelberg dual-channel supply chain. Omega 40(5):571–583
6. Chen W, Kucukyazici B et al (2016) Supply chain design for unlocking the value of remanufacturing under uncertainty. Eur J Oper Res 247(3):804–819
7. Dye CY, Yang CT (2016) Optimal dynamic pricing and preservation technology investment for deteriorating products with reference price effects. Omega 62:52–67
8. Hagiu A, Spulber D (2013) First-party content and coordination in two-sided markets. Manag Sci 59(4):933–949
9. Leider S (2016) Bargaining in supply chains. Manag Sci 62:3039–3058
10. Liu X (2016) Contracting for competitive supply chains under network externalities and demand uncertainty. Discrete Dyn Nat Soc 2016(1):1–9
11. Mantena R, Saha RL (2012) Co-opetition between differentiated platforms in two-sided markets. J Manag Inf Syst 29(2):109–140

12. Panda S, Modak NM et al (2015) Channel coordination and profit distribution in a social responsible three-layer supply chain. Int J Prod Econ 168:224–233
13. Shamir N (2015) Pricing strategy in dual channel under the leadership of retailers. Chin J Manag Sci 6(6):91–98
14. Shamir N (2016) Public forecast information sharing in a market with competing supply chains. Manag Sci J Inst Oper Res Manag Sci 62:2994–3022
15. Xie Y, Tai AH et al (2016) Pricing strategy for a two-echelon supply chain with optimized return effort level. Int J Prod Econ 182:185–195
16. Yang H, Zhuo W et al (2016) Two-period supply chain with flexible trade credit contract. Expert Syst Appl 66:95–105
17. Yao Z, Xu X, Luan J (2016) Impact of the downside risk of retailer on the supply chain coordination. Comput Ind Eng 12(102):340–350

Construction of Restoration System for Old Books Written in Braille

Yuko Shimomura[1], Hiroyuki Kawabe[1], Hidetaka Nambo[2(✉)], and Shuichi Seto[3]

[1] Kinjo University, Ishikawa, Japan
[2] Kanazawa University, Kanazawa, Japan
nambo@blitz.ec.t.kanazawa-u.ac.jp
[3] Kinjo College, Ishikawa, Japan

Abstract. Braille books follow a process of deterioration different from ordinary printed books. Braille is a reading letter and visually impaired people touch the boss with the finger's belly. Therefore, in Braille books that are frequently read, Braille is dirty with aged use, holes open, and they collapse. In old Braille books, braille collapses and deteriorates due to pressure from left and right books in the library. Therefore, in this research, we convert Braille books into machine-readable electronic data. First, a Braille book is scanned by an image. Next, we detect braille by image recognition technology one by one from braille page. And we classify and identify Braille. Furthermore, we correct errors such as misdetection of braille and misidentification. Finally, we save the result as character code.

Keywords: Braille · Old boo · Restoration · Deep learning

1 Introduction

In this section we briefly explain Braille and degraded braille.

1.1 Braille and Braille Library

The origin of Braille is old. In 1825, Braille [3] invented Braille in France. In 1879, Megata introduced Braille to Japan. In 1890, Ishikawa devised Japanese braille. This is the beginning of Japanese Braille. Since then Braille has been used as a character for the blind.

Braille consists of six embossed or flat points. These six points constitute one cell with two by three shape, which represents one character or a modifier of 63 kinds. With the modifier, characters more than 64 kinds are represented. In Japanese braille, each of voiceless sounds is represented by a plain cell and other characters such as syllabic nasals, voiced consonants, semi- voiced consonants, numerals, alphabets and so on are represented by a modifier and subsequent cell(s). In Fig. 1, Japanese Braille has a blank space for each clause like English.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_37

Some Japanese Braille are constructed with a prefix as last three examples in Fig. 1. Though Japanese is vertical writing or horizontal one, Japanese braille is horizontal writing only.

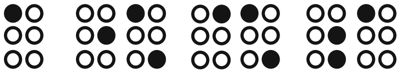


Fig. 1. Japanese Braille

The most primitive tool for writing braille is the slate and stylus. To use it, the user presses the tip of the stylus down through the small rectangular hole to make braille dots. Therefore the user is required to punch the braille dots with mirror image in reverse order. Nowadays, for publishing a book in braille, texts are input according to the braille grammar by using a braille text editor, are revised, and are printed in braille, as Fig. 2.

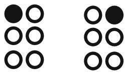


Fig. 2. Normal dot and mirrored one

In 1988, IBM opened "Tenyaku Hiroba (Field of Braille Transcription)", that was the Braille information network system, and digitization of braille had become widespread. The books in braille before the IBM's project are stored as they are without being converted into electronic data. Even if they have high storage value, the braille books are bulky, so braille libraries in various places are struggling to save. Each library has been forced to decide whether to discard them in recent years or to leave them.

In the earlier days when the Braille code was still in the research phase and there was no Windows, no personal computer, no internet, Shimomura et al. [1] studied coding of Braille.

Braille books are stored in the braille libraries throughout the country. The braille libraries registered in National Association of Institutions of Information Service for Visually Impaired Persons have about 100 libraries. Braille books. Braille books are borrowed from the braille libraries and public libraries via the braille libraries. In addition, many braille books are provided from the WEB braille library.

1.2 Degraded Braille

Braille is a tactile reading system. One reads tangible points with a finger's belly. Figure 3 is a typical page of Braille book. The braille books that are frequently

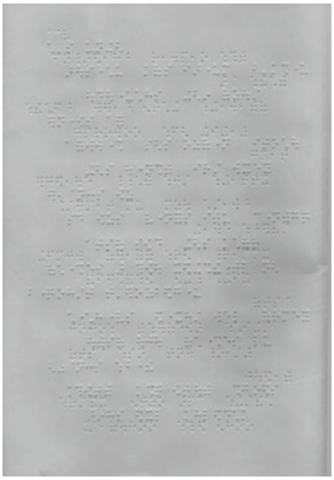


Fig. 3. Page of Braille book

read are dirty, holes are opened in dots, and dots collapse. Figures 4 and 5 show fresh dots and a part of a page that both normal cells and mirror images of cells, respectively. Figures 6, 7, 8 and 9 show the degraded braille cells, where the collapsed cell, the hole opened one, the dirty one and the distorted one is presented, respectively. Tears, creases and stains in pages are shown in Figs. 10, 11 and 12, respectively.

Fig. 4. Cell

Fig. 5. Normal and mirrored cell

Shimomura et al. [2] also studied restoration of the degraded Braille by shadow of Braille using fuzzy theory. In that study, the extraction of Braille was made by hand and computer programs executed the determination of existence of Braille. In this project, we extract Braille from scanned page images, which contain the degraded Braille, by computer programs, recognize them and restore the pages.

Fig. 6. Collapsed cell

Fig. 7. Hole opened cell

Fig. 8. Dirty cell

Fig. 9. Distorted cell

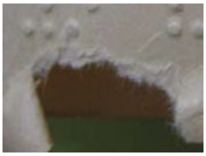


Fig. 10. Tears

Fig. 11. Creases

Fig. 12. Stains

2 Restoration System

In this subsection we explain our project of the restoration system for old books written in Braille.

(1) Page scanning.
(2) Determination of whether scanned dot is normal or mirrored.
(3) Extraction of cells and classification into 63 categories.
(4) Error correction by using a scanning redundancy, that one cell is read twice as the normal image and the mirror image.
(5) Error correction by using Braille grammar.
(6) Interpretation Braille into Japanese.
(7) Error correction by using Japanese grammar.

As noted above, the system has two main components, a machine learning system for recognition of Braille and an error correction system.

Fig. 13. Normal dot image

Fig. 14. Mirrored dot image

Fig. 15. Background image

2.1 Preparation of Braille Images for Machine Learning

To detect degrade cells by machine, we must prepare many image of degrade cells. We have already scanned many the old braille books with resolution of 200 dpi, extracted each cell into 54 $\times$ 36 pixel by hand, and have obtained about 15000 cell images. These images were classified into 63 normal categories and 63 mirror ones. Next, we cut each call images into 6 dot areas by a tiny application program and got about 45000 dot images and 45000 flat (non dot) ones. The followings are the normal image, the mirrored dot one and the background one for the machine learning (Figs. 13, 14 and 15).

2.2 Dot Recognition by OpenCV

OpenCV is a famous computer vision library and has functions of object detection. Some detectors such as face, eye, mouth and so on are prepared, and one also can make detectors for arbitrary objects. To make the object detector, one prepares many images one wants to detect and execute the machine learning. In the machine learning, characteristics are extracted from the many images of the object, and the machine learns the characteristics. The set of images and characteristics learned by the machine is called a cascade classifier. OpenCV has some algorithm to make the classifier. In our system, we adopt train the cascade routine with LBP characteristic to search dots in scanned page.

As the cells are regularly aligned on paper, by using the gradient of dots or the edge of paper we can correct the gradient of page. The distribution of detected dots determines row lines in page.

2.3 Scanning of Old Braille Books

It is necessary to treat the old Braille books carefully because of the degradation of cells in them as mentioned above. Therefore we do not adopt flatbed scanners but adopt a noncontact type of scanner, Fujitsu ScanSnap SV600, which operates by taking an elevated view (see Fig. 16), with resolution of 600 dpi and grey scale.

If a Braille book is printed on both side and all pages are scanned, all Braille printed is scanned twice. They are read as the normal image and the mirror image. This gives us a redundancy for interpretation of cells.

The next step is the dot detection by OpenCV mentioned above. The result is given in Figs. 17 and 18.

Fig. 16. ScanSnap SV600

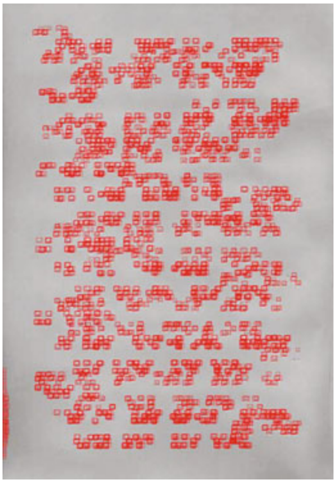


Fig. 17. Dot detection by OpenCV

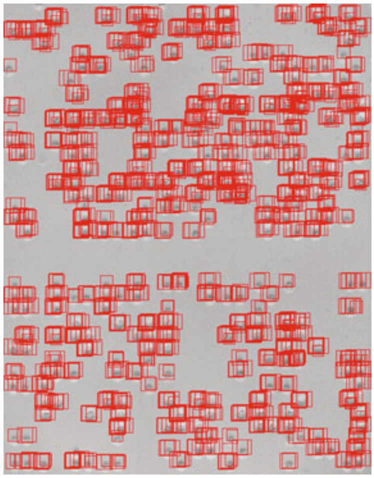


Fig. 18. Close up of detected dots

2.4 Normal and Reverse Dot Classification by Deep Neural Network

To detect degrade cells by machine, we must prepare many image of degrade cells. We have already scanned many the old braille books with resolution of 200 dpi, extracted each cell into 54 × 36 pixel by hand, and have obtained about 15000 cell images. These images were classified into 63 normal categories and 63 mirror ones. Next, we cut each call images into 6 dot areas by a tiny application program and got about 45000 dot images and 45000 flat (non dot) ones.

The deep learning is the one of technique of machine learning based on a deeply structured and hierarchical neural network. This technique is applied on the image recognition, the sound recognition and so on. In our project, we use this technique.

As usual Braille books are printed on both side, we must distinguish between the normal dots and the mirror ones. The both cell image are give again in Figs. 19 and 20. We have already obtained many dot images for them, so we use the images as reference data for the dot classification in the deep neural network.

This procedure gives us a normal dot distribution and a mirrored distribution. By converting the mirrored distribution into left-side right, we obtain the second candidate of the normal dot distribution.

Fig. 19. Normal cell

Fig. 20. Mirror cell

2.5 Cell Classification

As Braille is regularly aligned, the coordinate of dots tells us the starting point of cells. From this point, the dot images are sequentially analyzed, recognized and classified as cells by the deep neural network learned with 15000 cell images. Figures 20, 21, 22, 23, 24, 25 and 26 are binary aligned cell images. According to the classification, a character code is assigned to the cell image and is stored into database with a page number, a line number in the page, a word number in the line and a character number in the word. If the cell image has no character code, a flag of scanning error is also stored in the database.

2.6 Error Correction by Scanning Redundancy

When a Braille book is printed on both side and all pages are scanned, all pages become to be scanned twice. The character codes obtained from the normal images and ones from mirrored images are stored in the database and the codes obtained from normal and mirror images, for examples both shown in Figs. 16 and 17, are identical each other. By collating these codes and finding discrepancies, we have possibility to correct the scanning error.

2.7 Error Correction by Braille Grammar

Braille has own grammar. The character codes stored in the database are analyzed according to the grammar. In Japanese Braille, for example;

Fig. 21. "1xxxxx" cell

Fig. 22. "x2xxxx" cell

Fig. 23. "12xxxx" cell

Fig. 24. "1234 × 6" cell

Fig. 25. "12345x" cell

Fig. 26. "123456" cell

A postposition "ha" is translated into "wa".
There is not the postposition "ha" following numeric characters.
A long sound written with kana character is translated into a macron and so on. When unacceptable sequence of characters is found, a flag of Braille grammatical error is stored in the database. If a probable code can be presumed by the grammar, the code is also stored as a candidate in the database.

2.8 Error Correction by Japanese Grammar

According to Japanese grammar, we check the sequence of code, set flags of Japanese grammatical error for ungrammatical expressions and correct the codes to probable codes if possible. Typical grammatical errors are as follows;

- A sequence of punctuation marks;
- A sequence of contracted sounds;
- Non-correspondent parentheses Incorrect sonant marks.

2.9 Output and Correction by Hand

Finally the codes are converted into ink-spots expressing Braille and are output (See Fig. 27). For codes with the error flag, the ink-spots and candidates of Japanese character are output, and are corrected by hand.

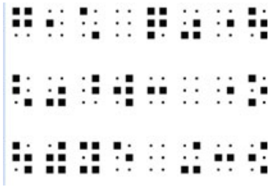


Fig. 27. Ink-spots expressing Braille

3 Concluding Remarks

In this paper, we explain our project to restore old Braille books by converting Braille books into machine-readable electronic data. The Braille book is scanned by an image. With the machine learning Braille are detected by image recognition technology, classified and identified. Furthermore, the error corrections are executed. Finally, the machine-readable character codes are stored.

Shimomura and colleagues have studied Braille 35 years before. The resumption of our study of Braille is because a request by the Japan Braille Library and Ishikawa Braille Library. In recent years, Braille books have been converted into electronic data and stored, and printed by computer processing. However the previous books are not converted into electronic data and are left as Braille books as it is. There are Braille books where the original is not found and without the original. The request was to convert Braille books into digital data before they became degraded and they became unreadable. Responding to this request, we resumed our research. We want to quickly restore degraded Braille books.

Acknowledgement. We would like to thank the staff of the following organizations for enabling us to research the actual condition of old Braille books: Ishikawa Association for Providing Facilities for the Visually Impaired and Japan Braille Library.

References

1. Shimomura Y, Mizuno S (1981) Translation of machine readable catalog/cataloging into braille and vice versa. Bull Kanazawa Women's Junior Coll 23:111–117
2. Shimomura Y, Mizuno S, Hasegawa S (1983) Automatic translation of machine readable data into braille. The Institute of Electronics, Information and Communication Engineers, IEICE Technical report ET82-9:57–62
3. National Association for Providing Facilities for the Visually Impaired (2003) Braille transcription guide, 3rd edn

Enterprise Operation Management

Discrimination of Personal Opinions from Commercial Speech Based on Subjectivity of Text

Yoshio Seino[1] and Takahiro Hayashi[2](✉)

[1] Graduate School of Science and Technology, Niigata University, 8050, Ikarashi-2-no-cho, Nishi-ku, Niigata-shi, Niigata 950-2181, Japan
[2] Institute of Science and Technology, Niigata University, 8050, Ikarashi-2-no-cho, Nishi-ku, Niigata-shi, Niigata 950-2181, Japan
hayashi@ie.niigata-u.ac.jp

Abstract. This paper presents a method for discriminating personal opinions from commercial speech. Today many personal opinions such as complaints about a particular product can be found on the Web. Mining useful information from these opinions is important for a wide range of applications. Personal opinions on the web, however, are often contaminated with commercial speech. Commercial speech is generated by companies and individuals for the intent of making a profit. The data cleaning process that discriminates personal opinions from commercial speech is important for obtaining useful results in opinion mining. As a data cleaning method, we propose a language-independent method for discriminating personal opinions from commercial speech based on subjectivity of text. Assuming that subjective words frequently occur in personal opinions rather than commercial speech, we define the subjectivity score of each word. Estimating the total subjectivity score of a given text, the proposed method identifies whether the given text expresses personal opinions or commercial comments. From experiments using texts datasets written in different languages, we have found over 90% of personal opinions can be correctly discriminated from commercial speech by the proposed method.

Keywords: Personal opinions · Commercial speech · Data mining · Data cleaning · Information filtering

1 Introduction

Gathering personal opinions such as complaints about a particular product is an important task for a wide range of applications. For example, in marketing, it helps identifying the potential revenues, judging the success of a new product launch, improving the quality of customer service, reducing the development cost of a next product, etc. Today, many personal opinions can be found on websites

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_38

of individuals, social networking services like Twitter and customer review pages in E-commerce sites. Owing to the importance of personal opinions on the Web, many opinion mining techniques have been proposed [3,5–9]. One of the major topics in opinion mining is to identify the polarity of a given text whether the expressed opinion in a text is positive, negative, or neutral.

However, texts collected from web sites are often contaminated with commercial speech. Commercial speech is generated by companies and individuals for the intent of making a profit, which is essentially different from personal opinions. Therefore, differentiating between personal opinions and commercial speech is important for obtaining useful results in opinion mining.

In our previous study, we proposed a method for discriminating personal opinions from commercial speech [1,2]. In the previous study, we assumed that subjective expressions such as negative meaning words like "bad" and emoticons like ":)" are more likely to be used for describing personal opinions than commercial speech. Under the assumption, we modeled the total subjectivity score of a given text for discriminating between personal opinions and commercial speech. However, in the previous method, the dictionary of subjective expressions was generated by human labor and it was strongly dependent on the Japanese grammar. It is, therefore, difficult to apply the method to other languages.

In this paper, we propose a language-independent method for discriminating personal opinions from commercial speech. The proposed method focuses on the statistical difference of appearance ratio of each word between personal opinions and commercial speech corpora. Based on the difference, the subjectivity of each word is estimated. Assuming that words having high subjectivity are frequently used in personal opinions, the proposed method judges whether a given text is personal opinions or commercial speech.

The rest of the paper is organized as follows. In Sect. 2, we explain the outline of related works. In Sect. 3, we describe the details of the proposed method. In Sect. 4, we explain experiments using corpora written in different languages. In the section we show that the proposed method can accurately discriminate the two types of documents independently from languages.

2 Related Works

The process of opinion mining can be divided into three phases: data collection, polarity analysis and post processing such as visualization of mining results. In this paper we focus on the data collection phase. In many studies on opinion mining aims at analyzing the polarity (positive, negative or neutral) of a given sentence or text [3,5–9]. Text corpora using in these studies are collected from various sources such as social networking service sites like Twitter and the customer review pages of E-commerce sites. These studies implicitly assume that all the analyzing documents are regarded as personal opinions. However, the texts gathered from these sources are often contaminated with commercial speech. There is obvious difference between minds of writers in personal opinions and commercial speech. Therefore, for people who want to collect real public impressions about subjects, personal opinions should be discriminated from commercial speech.

In our previous study [1,2], we proposed a method for discriminating personal opinions from commercial speech. In the previous method, in order to discriminate the two types of documents, we focused on the four kinds of subjective expressions: negative meaning expressions, sentence-final particles, interjections, and specific symbols such as emoticons. Measuring the appearance ratios of these expressions, we defined the subjectivity score of a given text. Using the scoring model, we judged texts having high subjectivity scores as personal opinions. However, the four kinds of subjective expressions were selected by human labor and these words are strongly dependent on the Japanese Grammar. Therefore, it is difficult to apply the previous method to other languages.

In this paper, we propose a language-independent method, where the dictionary of subjective expressions is automatically generated by analyzing two kinds of corpora: texts of personal opinions and ones of commercial speech. Based on statistical difference of appearance ratio of each word between the two corpora, the subjectivity score of each word is defined. Subjective words frequently occurring in personal opinions can be automatically detected. Since the scoring model is independently from languages, the proposed method can be applied to any languages by just chaining corpora.

3 Proposed Method

In order to discriminate personal opinions from commercial speech, the proposed method focuses on the difference of appearance ratios of words between two text corpora: corpus C_p and corpus C_n. Here, corpus C_p consists of text documents describing personal opinions and corpus C_n consists of text documents describing commercial speech. Using the two corpora, the subjectivity score of word w is defined as

$$s\left(w\right) = s_p\left(w\right) - s_n\left(w\right), \tag{1}$$

where $s_p(w)$ and $s_n(w)$ are defined as

$$s_p\left(w\right) = \frac{r_p\left(w\right)}{r_p\left(w\right) + r_n\left(w\right)}, \tag{2}$$

$$s_n\left(w\right) = \frac{r_n\left(w\right)}{r_p\left(w\right) + r_n\left(w\right)}, \tag{3}$$

here, $r_p(w)$ and $r_n(w)$ are the appearing ratios of word w in the corpus C_p and C_n, respectively. $r_p(w)$ and $r_n(w)$ are respectively defined as follows:

$$r_p\left(w\right) = \frac{\sum_{d \in C_p} f\left(w, d\right)}{\sum_{d \in C_p} \sum_{w' \in W(d)} f\left(w', d\right)}, \tag{4}$$

$$r_n\left(w\right) = \frac{\sum_{d \in C_n} f\left(w, d\right)}{\sum_{d \in C_n} \sum_{w' \in W(d)} f\left(w', d\right)}, \tag{5}$$

here, $W(d)$ is the set of words occur in document d, and $f(w, d)$ is the frequency of word w in document d.

Since subjective words like negative-meaning words are more frequently used in personal opinion rather than commercial speech [1,2], $s_p(w)$ is expected to be larger than $S_n(w)$ if word w is a subjective expression. The subjective scores of such words are expected to be positive ($s(w) > 0$). On the other hand, the subjective scores of words used in commercial speech tend to be negative.

Eliminating extremely-common words from a document is an important pre-process for text mining because these words give little meaningful information. In order to detect such common words, the document frequency of each word in a corpus is calculated. If word w occurs in over 99% documents in both corpus C_p and corpus C_n, the frequencies $f_p(w)$ and $f_n(w)$ are set as 0s.

Based on the subjective scores of words, the total subjectivity of a document d can be calculated as follows:

$$S\left(d\right) = \frac{1}{\left|W\left(d\right)\right|} \sum_{w \in W(d)} s\left(w\right), \tag{6}$$

where $W(d)$ is the set of words which appear in document d. If the subjectivity of a document is positive ($S(d) > 0$), the proposed method judges the text as personal opinions; if otherwise ($S(d) \leq 0$), the method judges the text as commercial speech.

4 Experimental Results

In order to confirm that the proposed method works independently from languages and subjects, we conducted 6 kinds of tasks (T1–T6) using different corpora. As corpus C_p (personal opinions) and corpus C_n (commercial speech) used in each task, we collected texts from different websites. Table 1 shows the source websites, the number of documents in each corpus, the language of the texts, and the topic of the texts. As shown in the table, in task T2, T3 and T4, the texts in corpus C_p and corpus C_n (personal opinions and commercial speech) were collected from the same websites. In these websites, documents are categorized into user reviews and professional reviews. In these tasks, the user reviews were assigned as the texts of corpus C_p (personal opinions) and the professional reviews were assigned as the texts of corpus C_n (commercial speech). There is obvious difference between minds of writers of the two types of documents. Professional writers tend to emphasize positive comments about subjects for their sponsors and seldom report negative comments about them. On the other hand, personal writers report any comments about subjects. For people who want to collect real public impression about subjects, user reviews would be more useful than professional comments. Therefore, professional reviews were regarded as commercial speech in the tasks.

The purpose of the experiments is to confirm whether the texts in corpus C_p (personal opinions) and the text in corpus C_n (commercial speech) were

Table 1. The details of corpora used in each task

Task	Language	Topic	♯ of documents	Sources of corpora
T1	Japanese	Product	C_p: 1980	Amazon.co.jp (https://www.amazon.co.jp/)
			C_n: 1975	Impress Watch (http://www.watch.impress.co.jp/)
T2	Japanese	Movie	C_p: 1496	映画.com (http://eiga.com/)
			C_n: 1706	映画.com (http://eiga.com/)
T3	English	Product	C_p: 4900	CNET.com (https://www.cnet.com/)
			C_n: 3803	CNET.com (https://www.cnet.com/)
T4	English	Movie	C_p: 990	Rotten Tomatoes (https://www.rottentomatoes.com/)
			C_n: 999	Rotten Tomatoes (https://www.rottentomatoes.com/)
T5	French	Product	C_p: 1147	Amazon.fr (https://www.amazon.fr/)
			C_n: 953	CNET France (http://www.cnetfrance.fr/)
T6	French	Movie	C_p: 692	SensCritique (https://www.senscritique.com/)
			C_n: 3093	Critikat (http://www.critikat.com/)

correctly classified by the proposed method. In addition, in order to clarify how the classification with the proposed method will generalize to an independent dataset, we used the technique of 10-fold cross validation [4]. That is, the texts in each corpus are divided into 10 subsets. By using 9 subsets as a training set, the subjectivity score of each word is defined by Eq. (1). Based on the subjectivity scores obtained with the training set, the classification is conducted for the texts in the remaining one subset. The training and classification are repeated 10 times by changing the combination of subsets for the training and test sets. Finally, the average performance across the 10 rounds is calculated.

For each round, the performance of classification results is evaluated with the F-value. The F-value is defined as follows:

$$F = \frac{2PR}{P + R}. \tag{7}$$

The F-value is the harmonic mean of the precision P and the recall R. The precision P and the recall R are defined as

$$P = \frac{|D_{\text{relevant}} \cap D_{\text{classified}}|}{|D_{\text{classified}}|}, \tag{8}$$

$$R = \frac{|D_{\text{relevant}} \cap D_{\text{classified}}|}{|D_{\text{relevant}}|}, \tag{9}$$

where D_{relevant} is the set of relevant texts, i.e., the texts of personal opinions in the test set, and $D_{\text{classified}}$ is the set of documents classified as personal opinions by the proposed method.

Table 2. The performances (F-values) of the proposed method in each task

Task	Maximum	Minimum	Average
T1	0.95	0.93	0.94
T2	0.93	0.88	0.90
T3	0.99	0.98	0.98
T4	0.99	0.97	0.98
T5	1.00	0.99	0.99
T6	0.98	0.96	0.97

The experimental results are shown in Table 2. In the table, the minimal, maximal and average F-values across the 10 rounds are shown for each task. From the table, we can confirm that in all the tasks (T1–T6), the high classification performances ($\bar{F} \geq 0.90$) were obtained. These results indicate that the proposed method can discriminate personal opinions from commercial speech independently from languages and subjects.

Figure 1 shows the distributions of subjectivity scores of the texts in corpus C_p and corpus C_n in used in each task. As shown in the figure, personal opinions and commercial speech can be clearly separated by the subjectivity score. These results indicate the effectiveness of the proposed scoring.

We analyzed frequently occurring words in personal opinions and commercial speech, respectively. Table 1(a) and (b) shows typical words in personal opinions and commercial speech, respectively. We have found that the words listed in Table 1(a) have high subjectivity scores and the words in Table 1(b) have negative subjectivity scores. As shown in the table, in personal opinions, subjective expressions, sentence-final particles (in Japanese), emoticons (in Japanese), and first-person singulars (in English and French) frequently occur. On the other hand, in commercial speech first-person plurals (in English and French) and neutral expression such as words just explaining the specification of a product. While negative-meaning expressions such as “bad” frequently occur in personal opinions regardless of its language, the appearance ratios of these words are

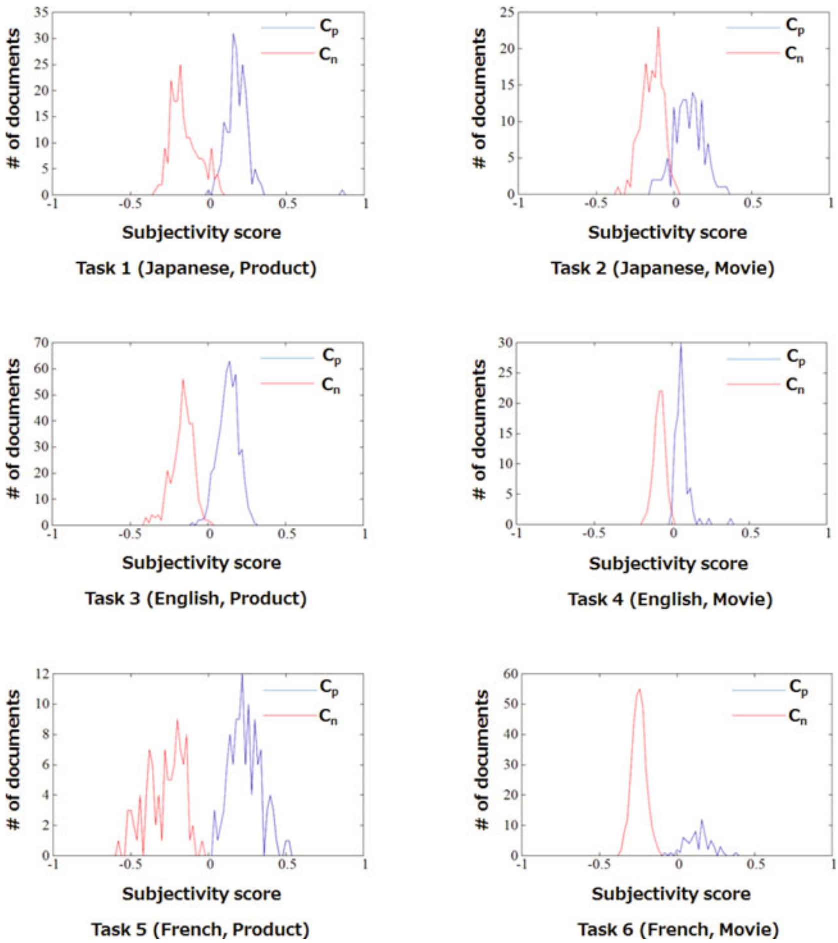


Fig. 1. The distributions of the subjectivity scores of the texts in corpus C_p (personal opinions) and corpus C_n (commercial speech) used each task

extremely lower in commercial speech. Sentence-final particles occur at the end of a sentence expressing linguistic modality. We have found sentence-final particles have high subjectivity scores, which indicate these words are useful for differentiating between personal opinions and commercial speech. Sentence-final particles are used in Asian languages including Japanese and Chinese. Further experiments applying the proposed method to other languages including Chinese texts remain as a future work (Table 3).

In the Indo-European languages including English and French, the subjects of sentences are helpful for differentiating personal opinions and commercial speech. For example, first-person singular is used in personal opinions but first-person plural is often used in commercial speech. These characteristics cannot be found

Table 3. Words frequently occurring in personal opinions and commercial speech

	(a) Personal opinions	(b) Commercial Speech
Japanese	**Subjective expressions** 悪い、良い **Sentence-final particles** かな、か、だな **Emoticons** (^^)/	**Neutral expressions** 液晶、搭載、監督、脚本
English	**Subjective expressions** good, bad, too **First person singular** I, my	**Neutral expressions** model, speed, drama, comedy **First person plural** We, our **Second person plural** You, your
French	**Subjective expressions** parfait, bon **First person singular** Je	**Neutral expressions** systeme, fiction **First person plural** Nous **Second person plural** Vous

in Japanese. This is because in Japanese the subject of a sentence is often absent. Although there is different usage of words in each language, with the proposed scoring method, informative words helping for decimation of personal opinions and commercial speech can be found automatically regardless of languages.

5 Conclusion

In this paper, we have presented a language-independent method for discriminating personal opinions from commercial speech. From the results of experiments using corpora written in different languages, we have confirmed that with the proposed method personal opinions can be effectively separated from commercial speech. Further experiments using other languages remain as a future work.

Opinion mining contributes to improve various phases in product-life cycle management such as improving the quality of customer service, creating new products, identifying the potential revenues, reducing the development cost, etc. The proposed method can be used as a method for data cleaning processing of opinion mining. Combining the proposed method with other post processing techniques in opinion mining such as sentiment analysis and visualization, we are planning to develop a comprehensive framework for opinion mining.

References

1. Hayashi T, Abe K, Onai R (2008) Retrieval of personal web documents by extracting subjective expressions. In: International conference on advanced information networking and applications - workshops, pp 1187–1192
2. Hayashi T, Abe K, Roy D, Onai R (2009) Discrimination of personal web pages by extracting subjective expressions. Int J Bus Intell Data Min 4(1):62–77
3. Khan FH, Bashir S, Qamar U (2014) Tom: Twitter opinion mining framework using hybrid classification scheme. Decis Support Syst 57(3):245–257
4. Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: International joint conference on artificial intelligence, pp 1137–1143
5. Rao S (2016) A survey on sentiment analysis and opinion mining. In: International conference on advances in information communication technology & computing, p 53
6. Ravi K, Ravi V (2015) A survey of opinion mining and sentiment analysis. J Knowl Based Syst 89(C):14–46
7. Selvam B, Abirami S (2013) A survey on opinion mining framework. Int J Adv Res Comput Commun Eng 3(9):3544–3549
8. Vinodhini G, Chandrasekaran RM (2012) Sentiment analysis and opinion mining: a survey. Int J Adv Res Comput Sci Softw Eng 2:282–292
9. You Q (2016) Sentiment and emotion analysis for social multimedia: Methodologies and applications. In: ACM on multimedia conference, pp 1445–1449

Shift in the Regional Balance of Power from Europe to Asia: A Case Study of ICT Industry

Zahid Latif[1(✉)], Jianqiu Zeng[1], Shafaq Salam[1], Zulfiqar Hussain[1], Lei Wang[1], Nasir Jan[2], and Muhammad Salman[3]

[1] School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
zahid25latif@yahoo.com
[2] School of Economics, Beijing Normal University, Beijing 100088, People's Republic of China
[3] Department of Economics, Gomal University, Khyber Pakhtunkhwa, Pakistan

Abstract. During the last couple of decades, ICT sector became the most innovative service sector that affected the living standards of human beings all over the world. In the beginning of the 21st century, some of the Asian countries made reforms in the ICT sector and spent an enormous amount for the progress of this sector. On the other hand, developed countries in the European Union (EU) faced different crises which badly affected the dissemination of this sector. Consequently, EU countries lost their hegemony in the field of information technology and resultantly, some of the emerging Asian countries like China, India, and South Korea got supremacy over the EU in this field. Currently, these countries have a strong IT infrastructure, R&D sector, IT research centers working for the development of ICT. Moreover, this paper investigates reasons for the shifting of the balance of digital power from Europe to Asia.

Keywords: Information and communication technology (ICT) · Infrastructure · R&D · One belt · One road · Game changer · Supremacy

1 Introduction

The global economy is facing many challenges as advanced economies are declining and emerging economies continue to expand their grounds. The level and duration of these developments are unprecedented. Also, recent developments in the field of information and telecommunication technology (ICT) such as cloud computing affects the trend of economies. We can define the trend as a shift in the regional balance of power. As the world's economic balance of power is shifting quickly, this trend has been accelerated due to the global recession. China overtakes the United States as the world's largest economic power and India enters the race of these two countries as another big economy. After the great

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_39

global recession, the world's balance of economic power in term of gross domestic product (GDP) was gradually shifted to the South and East [9]. Currently, the western industrialized countries are trying to get the growth momentum after the recession tragedy but they haven't fully recovered yet. On the other hand, the developing nations including Asian countries suffered comparatively less than the industrial countries and recovered rapidly after the recession. In the next years, the developing Asian countries accelerated their industrial growth, particularly in information and communication sector [12,25].

A steep decline has seen with the occurrence of 9/11, which resulted in the steady attrition of economic certainty that finally concluded with the great geopolitical obstruct of the western financial crisis [1,7]. It has now become the truth of the 21st century that the western world (Europe) is rapidly losing its superiority in the field of ICT and is replaced by a new international system shaped by geographical entity known as Asia [7]. Asia has the greater most number of internet users around the world and it covers almost half of the whole world. On the basis of internet usage and population statistics, Asia covers 45.7% of the world internet users, while rest covers the remaining 54.3% (See Table 1 and Fig. 1).

Table 1. Analysis of internet usage in Asia and rest of the world

Region	Population Est. 2016	% Population of the world	Internet users on 30/06/2016	Penetration (% Population)	% Internet users
Asia	4,052,652,889	55.20%	1,846,212,654	45.60%	50.20%
Rest of world	3,287,506,603	44.80%	1,829,612,159	55.70%	49.80%
World total	7,182,406,565	100%	3,035,749,340	42.30%	100%

Source: Internet World Stats, (www.internetworldstats.com/stats.htm)

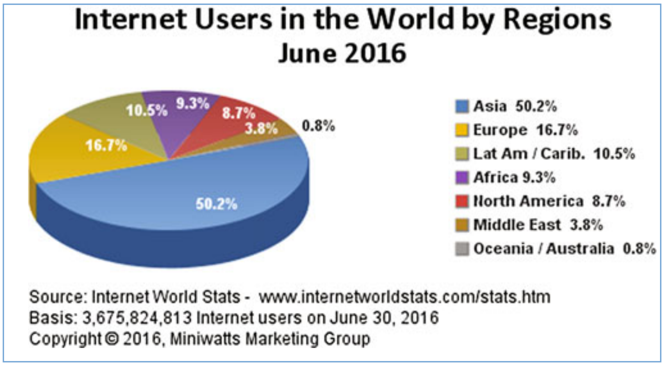


Fig. 1. Internet users in the world by regions

The above analysis shows the internet penetration and usage in Asia and rest of the world. The figures clearly describe the shift of technology towards Asia. Although due to low GDP and poverty, many people in Asian countries do not have access towards internet but the figures show richer tendency for these countries. China has one of the biggest communication network in Asia, as well as in the whole world. The ICT semiconductor industry has become one of the key industry of People Republic of China and has a huge potential to do business in ICT. This industry has undergone a rapid change of growth and development during the past decade [15]. The ICT sector in Asia includes goods and services that process, transmit or receive information. It includes technologies such as hardware, software, computer services, microelectronics, e-learning, e-business, e-health, multimedia, as well as emerging technologies such as photonics, fixed and mobile network convergence, life sciences, environmental sciences, Internet of Things, mobile internet, cloud computing and digital imaging [26].

2 Research Background

2.1 Literature Review

According to realist approach, the potential for clashes among great powers is an old story in the international system [4,8,19]. European Commission research scholars in EU Monthly Magazine predicted about the multi-polar world in 2025. It described that it is likely for the world to become truly multi-polar and reflecting the new balance of power, and the loss of US leadership. It further predicted that if the US remains the first military power, the scientific and technological advancement of some Asian states in new irregular war tactics like cyber-war and cyber-attacks will weaken the hegemony of US in information and communication technology [4,6,11,23]. Some authors give much importance to ICT even four decades before and predicted that the rise of a post-industrial society would be information centered and marked by a shift from production to service job [18,27,31].

The world's economic balance of power is shifting rapidly from North to South, and the trend has been accelerated by the global recession [1,9,30]. Kenichi Ohmae developed a concept of a "Triad", he predicted that the world economy including the information technology will be led by the United States, Japan, and the European Union [13]. But in present circumstances, the concept of Kenichi seems to have become concealed by a new order consisting of China, United States, and India in the field of ICT. Uri Dadush presented a model in a chapter of book "Handbook of Emerging Economies", in which he argued that the shift of technological balance from North to South is increased after the global recession. He considered technology as the most important element than the other factors for economic growth of both groups [16,24].

2.2 Conceptual Framework of ICT

The conceptual framework can be observed from the Fig. 2 which classifies economic and social benefits of ICT as well as the poverty outcomes. It clarify the concept that how ICT can promote economic and social benefits in a society.

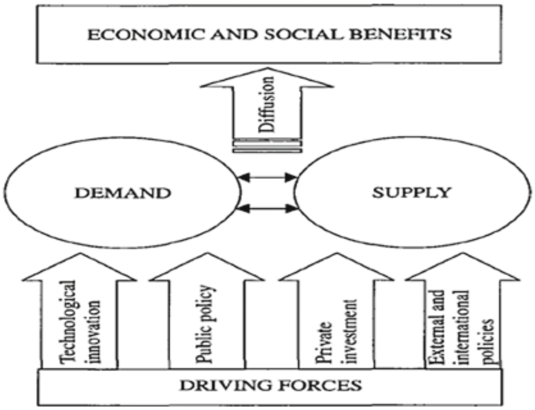


Fig. 2. A conceptual framework of ICT and development

The diffusion of ICTs, like other innovative technologies, proceed with three particular stages introduction, growth and maturity [17]. The ICT skills and infrastructure vary from country to country and is analyzed according to these three stages.

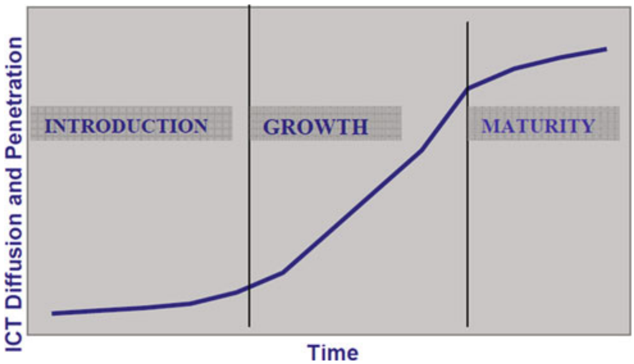


Fig. 3. ICT diffusion at three stages

Figure 3 shows that if a country developed, not only the mass of ICT users and practitioners increased but the necessary infrastructure for ICT must be developed. During the growth stage the diffusion rate is fast and is continuous until the maturity has been attained. The conceptual model of ICT diffusion will able the government and policy makers to adjust their goals and strategies in order to move from one stage to another stage. The proposed model will help the policy makers to improve the ICT skills and infrastructure in a way that it will reach to the next stages of the model.

3 A Comparative Analysis of Emerging Asian Economies and Europe

The recent development in ICT in the Asian region has accelerated the GDP growth of emerging Asian countries. Furthermore, the Asian region receives massive investments in the field of ICT, including both local and foreign investments. Most of the Asian countries attract massive investment through the sector of ICT. In this regard, the People Republic of China is on the top among the other emerging states followed by Russia, Brazil, and India. China is the biggest supplier of ICT equipment, while India has developed IT city called Silicon City (Bangalore).

The share of ICT in GDP has been decreasing gradually from the last several years in Europe, which dropped from 6.6% in 2009 to 5.9% in 2013. While the situation in emerging countries (mostly Asian and Asia pacific countries) in 2009 have accounted for third of the telecom services revenue and in 2013 it received almost 40% of the total world telecom revenue [3,5]. As a result, the gap in the growth rate and development of ICT between the developed countries and emerging countries became wider in 2013. In 2013, the balance in power regarding ICT kept on shifting to emerging countries continuously and become shared 80% of global growth contributing almost more than third of the world ICT market [29]. The comparison of ICT market growth and GDP are shown n the Table 2.

Table 2. Comparison of growth rates in advanced and emerging Asian countries

		2010	2011	2012	2013	2014
Advanced economies	Digiworld market growth	2.60%	1.30%	1.00%	1.10%	2.40%
	GDP growth	4.20%	3.70%	3.10%	2.60%	3.90%
Emerging Asian economies	Digiworld market growth	9.70%	10%	7%	7.70%	7.90%
	GDP growth	15.70%	13.20%	9.30%	10.10%	10.30%

4 Factors Influencing the Transfer of Digipower to Asia

The following are the some key reasons that cause the shift of balance of ICT in the favor of Asia from Europe.

- Single Telecommunication market (Monopoly)
- Fragmentation of digital market in Europe
- Influence of financial crisis in early years of 21st century
- Complex ICT Regulations that cause decrease in investment
- Economic Crisis
- Coalition against "war on terror"

Table 3. Global digiworld market by region

Billion ϵ	2012	2013	2014	2015	2018
North America	1073	1114	1166	1215	1349
Europe	924	925	929	955	1034
Asia/Pacific	993	1056	1123	1196	1411
Latin America	256	272	291	307	361
Africa/Middle-East	193	207	224	239	285
World	3439	3574	3732	3913	4439

The Table 3 shows the breakdown of digiworld in the region, it also have some predictions about the ICT market globally. The Fig. 4, also give a graphical picture of the percentage share of ICT market contribution throughout the world. The Figure clearly shows that the emerging Asian countries surpass the Europe in ICT market share following US.

4.1 Transfer of Digipower to Asia

Some scholars also state that the rise of Asia, particularly China will enable Beijing to pressurize US and the other western countries, not only through military power, but in the field of technology as well, which will lead to an "open and intense geopolitical rivalry". China will rise and become a new power throughout Asia [10,22,28]. According to the analysis of European Commission Joint Research Centre, China, Korea and Taiwan are the three highly specialized countries in ICT manufacturing, who have continued to strengthen their positions in the global ICT market [14]. China is the most obvious power on the rise in the present world scenario regarding ICT. In this study, China is the subject matter, because it not only belongs to the group of emerging economies growing very rapidly [2]. China took the lead among other regional states due to both its specialization in manufacturing of IT equipment and its economic size. Moreover,

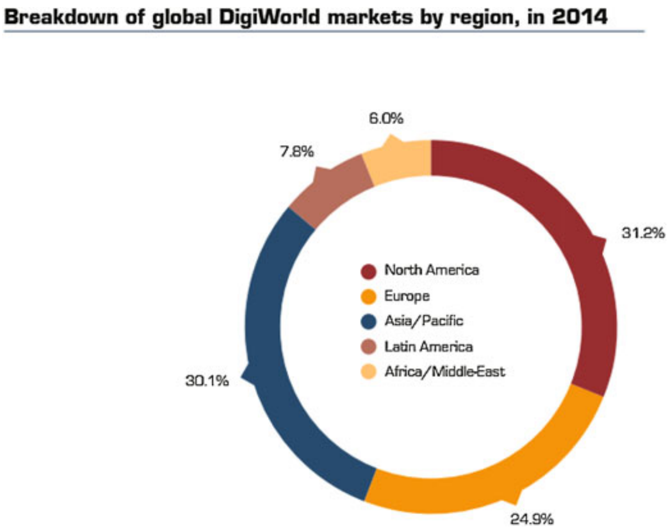


Fig. 4. Breakdown of global ICT market by region

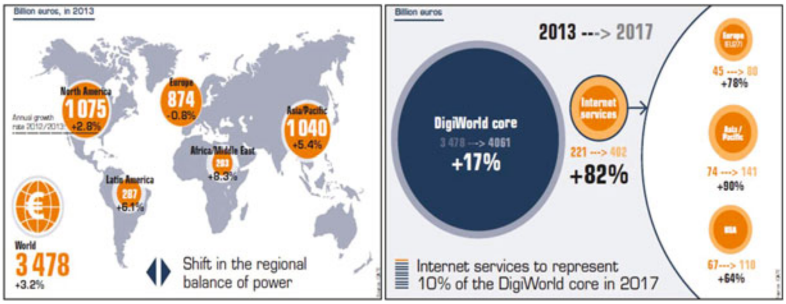


Fig. 5. Shift in the regional balance of power from Europe to Asia (ICT)

India and other emerging Asian states, which are now promoting their ICT sector, and can possibly surpass few major western states in the upcoming decades. The transfer of digital power ICT from West to Asia is getting velocity and may soon change the perspective for dealing with the global challenges. West is already aware of Asia's increasing potency in ICT.

The key lesson learnt from the international point of view is that, China, Korea and Taiwan having a large ICT manufacturing sector, have an important potential for growth, particularly, as the important resources like expenditures and qualified staff. In other words they have strengthened the ICT and R&D sector. While the US has dominance over EU due to its high productivity level of manufacturing as well as high R&D resources in order to retain its competitiveness in global ICT market [20, 21]. The detailed comparison and contribution of Asian countries and Europe in the world market (in ICT sector) are given in Fig. 5.

5 Conclusion

The facts and figures outlined in this paper show that the balance of power in information and communication technology (ICT) is drifting slowly and gradually towards Asia from Europe. Asia is a densely populated region with more than 2.8 billion people, living in a land fertile for IT and communication industry, with cheap labor and capital, rich raw resources and tremendous absorbing capacity for new technologies and developments. All these factors are molding the situation in favor of emerging Asian countries and shifting the regional balance of power in ICT towards Asia. The development of ICT has been the main objective of most of the Asian countries in changing the global environment towards the Asian region in terms of communication and technology. New Silk Road is an example of this strategy. Besides, US as the ICT giant has its own strategies towards occupying the information and communication market. This will require the sacrifice of core national interests in the short run. At the same time, China and other Asian emerging countries can get the dominancy in ICT sector only if these could ensure the interests of other countries according to their own ways in the long run.

References

1. Altman RC (2009) The great crash, 2008: a geopolitical setback for the west. Foreign Aff 88(1):2–0014
2. Bannister F, Connolly R (2014) ICT, public values and transformative government: a framework and programme for research. Gov Inf Q 31(1):119–128
3. Bastarrica F (2014) Mergers and acquisitions HR index. SSRN Electron J. http://dx.doi.org/10.2139/ssrn.2484775
4. Buzan B (2013) China and the US: comparable cases of 'peaceful rise'? Chin J Int Polit 6(2):109–132
5. Chesbrough HW (2006) The open innovation model: implications for innovation in Japan
6. Coen D (2009) Business lobbying in the European Union. Lobbying the European Union Institutions Actors & Issues
7. Cox M (2012) Power shifts, economic change and the decline of the west? Int Relat 26(4):369–388
8. Cox M, Booth K, Dunne T (1999) The interregnum: controversies in world politics 1989–1999. Cambridge University Press, Cambridge
9. Dadush U, Stancil B (2010) The world order in 2050
10. Friedberg AL (2000) Arming China against ourselves. Commentary 108(1):27–33
11. Halliday F (1996) Book review: Michael COX, US foreign policy after the cold war: superpower without a mission? Millenn J Int Stud 25(1):174–175 (London: Royal institute of international affairs and pinter, 1995, 148 pp, no price given)
12. Hawksworth J, Tiwari A (2011) The world in 2050: the accelerating shift of global economic power: challenges and opportunities. PwC-PricewaterhouseCoopers LLP, London
13. Hodara J, Ohmae K (1987) Beyond national borders: reflections on Japan and the world. Foreign Aff 65(5):1103

14. Hoge JF (2004) A global power shift in the making: is the united states ready? Foreign Aff 83(4):2–7
15. Kunigami A, Navas-Sabater J (2010) Options to increase access to telecommunications services in rural and low-income areas. Access & Download Statistics
16. Looney RE (2014) Handbook of emerging economies
17. Mamaghani F (2010) The social and economic impact of information and communication technology on developing countries: an analysis. Int J Manage 18(7):79–79
18. Masuda Y (1980) The information society: as post-industrial society. Institute for the Information Society
19. Mearsheimer JJ (2001) The tragedy of great power politics. Foreign Aff 80(6):173
20. Nepelski D, Prato GD (2012) Internationalisation of ICT R&D: a comparative analysis of Asia, the European Union, Japan, United States and the rest of the world. Asian J Technol Innov 20(2):219–238
21. Nepelski D, De Prato G, Stancik J (2011) Internationalisation of ICT R&D. Institute for Prospective Technological Studies, Joint Research Centre, European Commission
22. Ross RS (2006) Balance of power politics and the rise of China: accommodation and balancing in East Asia. Secur Stud 15(3):355–395
23. Ross RS (2012) The geography of the peace: East Asia in the twenty-first century. Int Secur 23(4):81–118
24. Schaub M (2009) Foreign investment in China-entry, operation and exit strategy. CCH Hong Kong Limited
25. Singh A (1997) Financial liberalisation, stockmarkets and economic development. Econ J 107(442):771–782
26. Sudan R, Bank W (2010) The global opportunity in it-based services: assessing and enhancing country competitiveness. World Bank Publications, pp 214–219
27. Toffler A (1991) Power shift: Knowledge, wealth, and violence at the edge of the 21st century. Powershift Knowl Wealth Violence Edge Century 5(4):91–92
28. Williams F (1989) Measuring the information society 4(1):82–83
29. Wohlers M, Giansante M et al. (2014) Shedding light on net neutrality: towards possible solutions for the Brazilian case. In: International Telecommunications Society Conference, pp 87–99
30. WorldBank (1997) Global economic prospects and the developing countries. WorldBank
31. Zhao Y (2014) Communication, crisis, & global power shifts: an introduction. Int J Commun 8(2):26

The Empirical Evidence of the Effect on the Enterprises R&D from Government Subsidies, Political Connections and Rent-Seeking

Dongliang Cai[1(✉)], Zhen Yang[2], Wu Jiang[1], and Qiuhua Xu[1]

[1] School of Finance, Southwestern University of Finance and Economics, Chengdu 610074, People's Republic of China
caidl@swufe.edu.cn

[2] Department of Finance, Hanqing Advanced Institute of Economics and Finance, Renmin University of China, Beijing 100872, People's Republic of China

Abstract. This study investigates the impact of rent-seeking behavior, political connections, and government science and technology (S&T) subsidies on the research and development (R&D) of private listed enterprises. Based on empirical data of China's private listed enterprises between 2008 and 2012, this paper performs validation at three levels and obtains suggestive conclusions. First of all, the purpose to obtain more S&T subsidies motivates private listed enterprises to take rent-seeking activities which can facilitate them in establishing political connections (PCs). Secondly, rent-seeking activities could eventually result in crowing-out effect of the R&D input, especially for the local PCs. Thirdly, the side effect of rent-seeking is more significant for enterprises located in provinces of high corruption levels, which weakens the availability of government S&T subsidies. The improper intervention of government distorts the market resources allocation.

Keywords: Rent-seeking · Political connections · Government subsidy · Research and development

1 Introduction

During the "supply-side reform", China's economic development mode is transforming from "factor-driven" to "innovation-driven", so it is critical to strengthen the R&D of enterprise innovations. But funding and policy restraints make it difficult for small and medium-size private enterprises to carry out project research. Since the process of research and innovation is a long high-risk investment period, it is necessary for governments to provide various incentives and subsidies so as to promote enterprises R&D. However, through the economic transition, private listed enterprises, especially small ones, suffered from

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_40

"explicit" or "hidden" discrimination in policies especially in terms of financial resources and government subsidies.

Plenty researches have recognized that enterprise can promote development by political connections, benefiting from the government subsidies, financing convenience, and tax preference [2,6,21]. Since PCs can bring huge economic benefits, enterprises are intent to ingratiate with government officials by rent-seeking activities for the purpose of getting more S&T subsidies [17]. There are many studies investigating the effects of the PCs and government subsidies on the enterprise performance. Wu [22] finds the state-owned enterprises can receive more subsidies than private enterprises, which points out that PCs can help enterprises obtain S&T subsidies. Meuleman and Maeseneire [16] found the government subsidies promote enterprises to innovate. Xu [25] analyze the signal effect of the government subsidies. But on the other side, S&T subsidies will crowd out the enterprise R&D input [26]. Some empirical results indicate that political connections are more likely to serve a binding function [28]. However, beyond that, it still lacks detailed discussions of how PCs and S&T subsidies affect the R&D input, fairly clear explanations of why the S&T subsidies failed to make effects, and intuitive evidences of the crowding-out effect of rent-seeking corruptions on R&D input.

In order to enrich the research of this field, this paper designs the three levels of verification: first, to test if rent-seeking activities help maintain the PCs; second to verify if it will crowd out the enterprise R&D input; lastly, to investigate if corruption levels in enterprise provinces will make a difference and whether it will ruin the utility of S&T subsidies.

The contributions of this study are as follows. Firstly, it provides empirical evidence of the economic consequences of establishing political connections and obtaining S&T subsidies on the enterprise R&D input. Secondly, it enriches the explanation for the failure of government subsidies. It is found that although rent-seeking activities facilitate establishing PCs and makes it easier to get government S&T subsidies, it reduces the utility of government subsidies. Thirdly, this study also provides microcosmic evidence that government S&T subsidies may distort the social resource allocation under certain conditions.

2 Theoretical Analysis and Research Hypothesis

2.1 Rent-Seeking and Political Connections

Political connections mean close connections between the internal personnel of the enterprise and government officials, or participation in political affairs. However, political connections do not mean corruption [6]. The existing financial and legal systems in China still cannot fully meet the various needs of the transition of economic development patterns, and the distribution of resources is mainly driven by the government, which all lead to the current situation that enterprises scramble for more opportunities by seeking for informal ways under the development pressure [1,15]. Under the special period of "adjusting structure and transition of patterns", and increasingly drastic market competition in China, enterprises are trying to gain more financial resources or policy welfares through such "recessive capitals", namely the PCs [23,29,30].

In these situations, enterprises in China must have strong motivation to establish PCs by rent-seeking [4]. Therefore, the following hypothesis is proposed:

Hypothesis 1: Rent-seeking can help enterprises establish PCs.

2.2 Rent-Seeking and Enterprise R&D

Now, the formal institutions are weak in China [5]. Corruption and rent-seeking distort fiscal and monetary policy [3,8]. Entrepreneurs are more likely to spend time and resources on rent seeking rather than on productive activities in order to influence governments [4]. Under weak institutional contexts, enterprises have to satisfy government's requirements, which aggravates burden on enterprises [14]. In this condition, government subsidies can deteriorate corporate performance. Quevedo [7] proposed that government S&T subsidies have a negative effect on the R&D expenditure, namely the crowding-out effect. Huang [9] made a case study on Chinese listed companies between 2001 and 2007, finding that PCs have a promoting effect on the performance of SOEs, but negative impact on private listed enterprises.

Following the normal logic of enterprise behavior, the following hypothesis is proposed:

Hypothesis 2: Enterprises are intent to establish PCs by rent-seeking for obtaining more government S&T subsidies, but it will lead to crowding-out effect on R&D input.

2.3 S&T Subsidies and Enterprise R&D

As is recognized, government S&T subsidy are important for enterprises, which can alleviate market failures in R&D activities [11,24]. Levin and Reiss [13] studied the relationship between government subsidies and R&D output on the micro level, finding that government subsidies have a remarkable role in promoting R&D. Lee and Cin [12] drew the same conclusions by considering different country samples and selecting panel data for several years. The government R&D subsidies contribute to both green innovations [19] and sustainable innovations [10]. Yet, there are opposite opinions. Zhang [27] found subsidies, in long and short terms, have significant positive effect on the financial performance of wind energy manufacturing companies. In renewable energies sector, subsidy policies could cause adverse consequences to a certain extent [20]. The current economic transition period in China, is combined with a serious rent-seeking phenomenon and ethical risks, so the establishment of PCs may consume a tremendous amount of manpower and material resources for enterprises and thus fail to improve the performance.

Therefore, the following hypothesis is proposed.

Hypothesis 3: In provinces with serious corruptions, the rent-seeking results in more significant crowding-out effect, frustrating the availability of government S&T subsidies.

3 Research Design and Sample Selection

3.1 Sample

The original sample contains the non-financial private A-share companies listed in Shanghai and Shenzhen Stock Exchange during 2008–2012. The final sample keeps 361 companies listed 1614 totally after basic sample processing.

The major data used in the empirical part includes: the detailed government subsidies data, PCs data, basic enterprise data, post crime data at the provincial level and the excess administration expense data. The data was derived from the CSMAR and WIND database, the China inspection yearbook and China statistical yearbook. Most of them need manual calculation.

3.2 Model Setup and Variable Definition

To test *Hypothesis 1*, we set up the following probit regression model

$$P_{\mathrm{Poli},i,t} = \alpha + \beta_1 \times \mathrm{Rent}_{i,t} + \beta_2 \times X_{i,t} + \varepsilon_{i,t}. \tag{1}$$

To illustrate the robustness of conclusions we also set up the following logistic model:

$$\log\left(\frac{P_{\mathrm{Poli},i,t}}{1 - P_{\mathrm{Poli},i,t}}\right) = \alpha + \beta_1 \times \mathrm{Rent}_{i,t} + \beta_2 \times X_{i,t} + \varepsilon_{i,t}. \tag{2}$$

Here, $P_{\mathrm{Poli},i,t}$ is the probability that company i established political connections in year t. Correspondingly, $\mathrm{Poli}_{i,t}$ represents the dummy variable of PCs in year t while Loca represents the dummy variable of central PCs. Then, referring to the treatment methods proposed by Yu et al. [17], the participation in the politics of private listed companies is depicted by whether the directors and senior executives of listed companies have held a position in the National People's Congress (NPC), Chinese People's Political Consultative Conference (CPPCC) or government sectors. Specifically, $Poli_{i,t}$ is defined as 1 if the directors or senior executives have been the members of the NPC or CPPCC, or held a position at a government sector, otherwise it is 0. $\mathrm{Loca}_{i,t}$ is the same way defined for the local NPC, CPPCC, or local government sectors.

$\mathrm{Rent}_{i,t}$ represents the excess administration expense of company, a proxy variable of rent-seeking, which should be calculated as with a regression process introduced by RichardsonM [18]. The excess administration expense is the in-sample forecast residuals of the regression between the ratio (management cost/operating income) with a series of factors based on the historical data.

To verify *Hypothesis 2*, we set up the following model:

$$\log \mathrm{Devlpexp}_{i,t} = \alpha + \beta_1 \times \log \mathrm{SciSubsiby}_{i,t} + \beta_2 \times \mathrm{Rent}_{i,t} + \beta_3 \times X_{i,t} + \varepsilon_{i,t}, \tag{3}$$

where $\log \mathrm{Devlpexp}$ is the logarithm of corporate R&D investment; $\log \mathrm{SciSubsidy}$ represents the enterprises' S&T subsidies from government.

To verify *Hypothesis 3*, we set up the following model:

$$\begin{aligned}\log \mathrm{Devlpexp}_{i,t} = \alpha &+ \beta_1 \times \log \mathrm{SciSubsiby}_{i,t} + \beta_2 \times \mathrm{Rent}_{i,t} \\ &+ \beta_3 \times Poli_{i,t} + \beta_4 \times X_{i,t} + \varepsilon_{i,t}.\end{aligned} \tag{4}$$

On the basis of model (3), model (4) adds the PCs variable Poli, and the sub-sample regression for local PCs companies. In addition, the regression is divided into different groups according to the corruption level in the provinces. The corruption level is depicted by the corruption and bribery cases registered every 10 thousand public officers, which is exactly the ratio of registered post criminal cases at the provincial level to the number of civil servants.

In the model (1) & (2), X is a multiple vector of control variables including: the logarithmic company size ($\log$ Asset), rate of asset (ROA), debt-to-assets ratio(Dbassrt), establishment years (EstAge), the logarithm of top three executive compensations (LogMane3Pay), the duality of CEO and chairman

Table 1. Descriptive statistics of variables from groups with different PCs (10,000 yuan)

	Companies without PCs			Companies with PCs			Companies with local PCs		
	Mean	sd	Med	Mean	sd	Med	Mean	sd	Med
S&T subsidies	2.62	6.52	0	4.54	31.21	0	4.75	33.4	0
Operating income	3037.4	3993.23	1607.68	6495.62	13721.8	2260.71	5423.21	12418.35	1929.38
R&D input	45.08	89.64	14.06	77.18	248.51	16.44	59.19	241.09	12.15
Overheads	187.82	224.04	119.54	298.49	554.39	136.26	247.64	451.23	114.78
Top three executive compensation	1.18	1.17	0.9	1.34	1.36	1.05	1.16	0.9	0.97
Excess administration expense	−0.01	0.12	−0.01	0	0.15	−0.01	0	0.07	−0.01
Top three directors' compensation	1	0.97	0.75	1.27	1.61	0.88	1.04	1.03	0.79
The proportion of independent directors	0.37	0.05	0.36	0.36	0.05	0.33	0.36	0.04	0.33
The duality of CEO and chairman	0.19	0.39	0	0.18	0.38	0	0.24	0.43	0
Corruption level	24.08	5.67	24.31	24.52	6.78	24.31	24.4	7.34	24.31

(Chairman&CEO is defined as 1 for duality), the proportion of independent directors (IndeDirRatio), year control variable, industry control variable (Industry is defined as 1 if the company belongs to the national key support, highly regulated or monopoly industries, otherwise 0 [17].

3.3 Descriptive Statistics

It can be seen from Table 1, politically connected companies obtain more government S&T subsidies in average which is consistent with the previous literatures view that political connections help obtain the government subsidies. At the same time, the politically connected enterprises have relatively higher overheads and excess administration expense, indicating that rent-seeking may exist.

Table 2. Descriptive statistics of variables between groups with different excess administration expense

	Low excess administration expense			High excess administration expense		
	Mean	sd	Med	Mean	sd	Med
S&T subsidies (10,000 yuan)	5.87	24.78	0	10.47	51.32	0
T&D input (10,000 yuan)	44.51	107.41	16.34	148.96	359.71	41.77
Overheads (10,000 yuan)	192.9	230.08	130.62	531.86	773.28	287.51
R&D input/Overheads	0.21	0.23	0.11	0.23	0.26	0.13
Operating income (10,000 yuan)	3903.8	6600.95	2171.7	11630.4	18952.16	4605.86
Overheads/ Operating income	0.07	0.04	0.06	0.13	0.27	0.07
Total asset (10,000 yuan)	5773.42	10141.05	2899.96	11278.23	17493.17	5611.22
Top three executive compensation (10,000 yuan)	1.18	0.89	0.96	1.54	1.83	1.14
Top three directors' compensation (10,000 yuan)	1.1	1.28	0.8	1.36	1.9	0.93
The proportion of independent directors	0.37	0.05	0.33	0.36	0.05	0.33
The duality of CEO and chairman	0.16	0.37	0	0.13	0.33	0

On the basis of Table 2, though the enterprises with high excess administration expense can obtain more S&T subsidies, they do not show any significantly higher ratio of R&D input/overheads but get higher ratio of the overheads/operating income. It may be caused by the crowding out effect from the rent-seeking behaviors, namely the existence of corruption and rent-seeking reduce the efficiency of the use of government S&T subsidies.

4 Empirical Results and Analysis

4.1 Rent-Seeking Behavior and Political Connections

Table 3 contents 4 regression results of the probit model and logistic model used to test the influence of excess administration expense on the total polit-

Table 3. Test results of the hypothesis 1

Variable	P_{Poli}		$\log(P_{\text{Poli}}/(1-P_{\text{Poli}}))$	
	Probit Model		Logistic Model	
Rent	0.319**	0.222*	0.541**	0.378*
	(0.127)	(0.128)	(0.224)	(0.225)
L1Rent	-	0.018*	-	0.031
	-	(0.010)	-	−0.021
logAsset	−0.088*	−0.104*	−0.158*	−0.187*
	(0.053)	(0.058)	(0.092)	(0.101)
ROA	−0.378	−0.468	−0.635	−0.756
	(0.25)	(0.299)	(0.394)	(0.506)
Dbassrt	−0.088	−0.23	−0.158	−0.411
	(0.169)	(0.19)	(0.28)	(0.314)
EstAge	0.018	0.014	0.032	0.026
	(0.012)	(0.014)	(0.022)	(0.026)
logMane3Pay	−0.151**	−0.126	−0.258**	−0.211
	(0.07)	(0.077)	(0.119)	(0.131)
Chairman& CEO	−0.239**	−0.259**	−0.392**	−0.415*
	(0.113)	(0.126)	(0.19)	(0.213)
IndeDirRatio	−2.529***	−1.539	−4.059**	−2.384
	(0.96)	(1.067)	(1.608)	(1.784)
_cons	5.337***	5.150***	9.115***	8.757***
	(1.425)	(1.538)	(2.45)	(2.649)
Year	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
R^2	0.052	0.055	0.052	0.054
N	925	727	925	727

Note: Robust standard errors in parentheses, ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

ical connections. The regression results of probit model reflect a significantly positive coefficients of Rent and the first-lagged Rent as well. It suggests that excess administration expense can help enterprises establish political connections, which is companied with hysteresis effect. To maintain the robustness of the conclusion, the logistic model is conducted at the same time and the results are basically the same as the probit models, which shows a robust estimating conclusion.

4.2 Rent-Seeking Behavior and Enterprise R&D Input

Based on the results of the *Hausman* test and *Wald* test, corresponding regression models are applied to test the influences of government S&T subsidies and excess administration expenses on enterprise R&D input. Table 4 contents the regression results of groups with different political connections. Comparing the regression in columns 1 and 2, the coefficient of *Rent* is significantly negative for companies with PCs, indicating that high excess administration expense will probably reduce the R&D input. Overall, test results of the Hypothesis 1 draw the conclusion that enterprises are intend to establish political connections through rent-seeking which makes it easier to acquire government S&T subsidy. According to the motivation of establishing PCs, we can infer that the pursuance of government S&T subsidies will stimulate more rent-seeking activities which eventually result in crowding-out effect on the primary business performance.

Under the fiscal decentralization system in China, local governments are empowered with great administrative power and relatively weak financial power since the tax reform. When local governments try to obtain more resources through extra-budgetary way with less supervision, it is prone to create rent-seeking corruptions between local governments and enterprises. At the same time, the local governmental officials have strong discretionary power in science subsidy and financial support (including rewards) which has no explicit restriction in laws or regulations. So the difference between enterprises with and without local political connections are further investigated as regression results in columns 3 and 4. The variable Rent has a significantly larger negative coefficient than total sample in columns 1, which suggests that local PCs will cause higher rent-seeking cost and worse impact on the enterprise's performance.

4.3 R&T Subsidies and Corporate R&D

Following the *Hausman* test and *LR* test, corresponding regression models are applied to test the hypothesis 3 and results are listed in Table 5. Compared to the low corruption group, high corruption group in columns 1 and 2 showed distinctly negative effect of variable Rent. The figures illustrate that rent-seeking behavior will lead to substantial crowding-out impact on the R&D of enterprises in provinces with serious corruptions. On the other hand, coefficients of log SciSubsidy are almost significantly negative in columns 1 and 2, but positive in columns 3 and 4. Thus for high corruption group, the government S&T subsidies probably cause negative impact on the enterprise R&D. It is because enterprises have to invest great manpower and resources for rent-seeking, which will fail the fulfill use of S&T subsidies and crowd out scientific research innovation input. In addition, the coefficient of Loca is significantly negative in column 2, suggesting that the local PCs will strengthen such side effect, which is consistent with the results in Table 4.

Table 4. Test results of the Hypothesis 2

Variable	log Delpexp			
	Companies with PCs	Companies without PCs	Companies with local PCs	Companies without local PCs
log SciSusidy	with PCs 0.073^{***} (0.027)	without PCs 0.063 (0.217)	local PCs 0.01 (0.047)	local PCs 0.075^{**} (0.032)
Rent	-3.427^{***} (1.057)	11.83 (7.729)	-5.131^{***} (1.182)	-3.757^{*} (2.178)
log Asset	0.485^{***} (0.119)	0.818 (2.199)	0.352 (0.229)	0.617^{***} (0.127)
ROA	4.009^{**} (1.636)	8.815 (8.911)	0.956 (3.319)	3.860^{**} (1.815)
Dbassrt	-1.718^{**} (0.725)	1.274 (3.023)	0.735 (1.213)	-2.201^{***} (0.782)
EstAge	0.004 (0.016)	0.674 (1.09)	−0.056 (0.062)	0.002 (0.016)
log Mane3Pay	0.457^{***} (0.163)	−0.459 (0.716)	0.658^{**} (0.301)	0.512^{***} (0.182)
ChairmanCEO	−0.105 (0.23)	-1.427^{***} (0.243)	0.704 (0.482)	0.356 (0.33)
IndeDirRatio	-5.961^{**} (2.89)	−0.782 (8.063)	3.087 (5.342)	-4.518^{*} (2.639)
cons	2.829 (3.674)	−8.068 (25.669)	0.121 (7.705)	−1.098 (3.363)
Year	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
r^2	0.371	0.785	0.526	0.395
N	192	49	64	177

Note: Robust standard errors in parentheses, ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Table 5. Test results of the hypothesis 3

Variable	logDelpexp			
	High corruption		Low corruption	
log SciSubsidy	−0.044 (0.032)	−0.119*** (0.046)	0.106** (0.058)	0.095* (0.066)
Rent	−6.180*** (0.818)	−5.567*** (1.051)	−0.020 (1.859)	−0.115 (2.053)
log Asset	0.595** (0.232)	0.507** (0.247)	0.611*** (0.204)	0.855*** (0.250)
ROA	3.687 (3.745)	−1.426 (2.402)	5.704* (3.713)	3.225 (3.424)
Dbassrt	−2.181 (1.741)	−1.514 (1.533)	−0.651 (1.207)	−0.966 (1.209)
EstAge	0.034 (0.054)	0.011 (0.066)	0.009 (0.018)	0.024 (0.024)
log Mane3Pay	0.994*** (0.324)	0.948*** (0.280)	0.159 (0.300)	0.084 (0.319)
ChairmanCEO	−0.364 (0.703)	0.204 (0.732)	0.000 (0.479)	−0.057 (0.409)
IndeDirRatio	1.163 (2.960)	−2.755 (2.324)	−4.325 (5.482)	−5.091 (5.076)
Poli	−0.175 (0.393)	- -	0.257 (0.379)	- -
Loca	- -	−1.136*** (0.430)	- -	0.412 (0.449)
cons	−8.135 (7.067)	−2.900 (7.119)	1.771 (5.826)	−2.137 (6.425)
Year	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
R^2	0.505	0.584	0.479	0.434
N	62	62	82	82

Note: Robust standard errors in parentheses, ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

4.4 Robustness Analysis

To guarantee the robustness of conclusions, we build logistic and probit models. Considering the possible endogenous problems, the lagged excess administration expense is added into the test of Hypothesis 1, As for hypothesis 2 and 3, we use the current and previous budget of technology spending from enterprise‘s local government as the instrumental variable to conduct two-stage least squares regressions, the exogeneity of the government budget avoid the two-way causality problem. The results are basically consistent, indicating the reliability of conclusions.

5 Research Conclusions

We find evidence that rent-seeking behavior facilitate private listed enterprises in establishing PCs and thus obtaining more S&T subsidies. However, it is proved that rent-seeking activities will finally result in crowing-out effect of the R&D input, especially for the enterprises located in the province of higher corruptions. For such companies, the PCs are closely related to the S&T subsidies, which breeds severe rent-seeking corruption and weakens the availability of subsidies. Our results support the idea that PCs can result in rent-seeking corruption, because enterprises are motivated to ingratiate with government officials by rent-seeking activities for the purpose of getting more S&T subsidies. And it turns out that the crowding-out effect of such activities will eventually fail the availability of subsidies. The PCs and S&T subsidies cannot effectively promote corporate performance or improve the development of the whole society, but only serve for a few entrepreneurs and politicians. The government should regulate the allocation of government subsidies, and have a reasonable control over local government powers to improve the efficiency of the whole market and stimulate the enthusiasm of enterprise innovation.

Acknowledgements. This research was supported by the project of Research Center for System Sciences and Enterprise Development (Grant No. Xq16C03).

References

1. Allen F, Qian J, Qian M (2005) Law, finance, and economic growth in china. J Financ Econ 77(1):57–116
2. Claessens S, Feijen E, Laeven L (2008) Political connections and preferential access to finance: the role of campaign contributions. J Financ Econ 88(3):554–580
3. Dimakou O (2009) Bureaucratic corruption and the dynamic interaction between monetary and fiscal policy. Eur J Polit Econ 40:57–78
4. Du J, Mickiewicz T (2015) Subsidies, rent seeking and performance: being young, small or private in China. J Bus Ventur 31(1):22–38
5. Estrin S, Korosteleva J, Mickiewicz T (2013) Which institutions encourage entrepreneurial growth aspirations? J Bus Ventur 28(4):564–580
6. Faccio M (2006) Politically connected firms. Am Econ Rev 96(1):369–386
7. Garcła-Quevedo J (2004) Do public subsidies complement business R&D? A meta-analysis of the econometric evidence. Kyklos 57(1):87–102
8. Hessami Z (2014) Political corruption, public procurement, and budget composition: theory and evidence from OECD countries. Eur J Polit Econ 34(6):372–389
9. Huang MC, Wu CC (2016) Facts or fates of investors' losses during crises? Evidence from reit-stock volatility and tail dependence structures. Int Rev Econ Financ 42:54–71
10. Ketata I, Sofka W, Grimpe C (2015) The role of internal capabilities and firms' environment for sustainable innovation: evidence for germany. R&D Manage 45(1):60–75
11. Kleer R (2010) Government R&D subsidies as a signal for private investors. Res Policy 39(10):1361–1374

12. Lee EY, Cin BC (2010) The effect of risk-sharing government subsidy on corporate R&D investment: empirical evidence from korea. Technol Forecast Soc Change 77(6):881–890
13. Levin RC, Reiss PC (1984) Tests of a schumpeterian model of R&D and market structure. University of Chicago Press, pp 175–208
14. Marquis C, Qian C (2013) Corporate social responsibility reporting in China: symbol or substance? Organ Sci 25(1):127–148
15. Mellahi K et al (2016) A review of the nonmarket strategy literature: toward a multi-theoretical integration. J Manage 42(1):143–173
16. Meuleman M, Maeseneire WD (2012) Do R&D subsidies affect SMEs' access to external financing? Soc Sci Electron Publishing 41(3):580–591
17. Minggui Y, Yafu H, Hongbo P (2010) Political connection, rent seeking and the effectiveness of fiscal subsidies by the local government. Econ Res J 3:65–77
18. Richardson S (2006) Over-investment of free cash flow. Rev Acc Stud 11(2):159–189
19. Schiederig T, Tietze F, Herstatt C (2012) Green innovation in technology and innovation management-an exploratory literature review. R&D Manage 42(2):180–192
20. Shen J, Luo C (2015) Overall review of renewable energy subsidy policies in China-contradictions of intentions and effects. Renew Sustain Energy Rev 41:1478–1488
21. Su ZQ, Fung HG et al (2014) Cash dividends, expropriation, and political connections: evidence from China. Int Rev Econ Financ 29(1):260–272
22. Wu A (2016) The signal effect of government R&D subsidies in China: do ownership matter? Technol Forecast Soc Change 117:339–345
23. Xu C (2011) The fundamental institutions of China's reforms and development. J Econ Lit 49(4):1076–1151
24. Xu E, Xu K (2013) A multilevel analysis of the effect of taxation incentives on innovation performance. IEEE Trans Eng Manage 60(1):137–147
25. Xu K, Huang KF, Xu E (2014) Giving fish or teaching to fish? An empirical study of the effects of government research and development policies. R&D Manage 44(5):484–497
26. Yu F, Guo Y et al (2016) The impact of government subsidies and enterprises' R&D investment: a panel data study from renewable energy in China. Energy Policy 89:106–113
27. Zhang H, Li L et al (2014) Political connections, government subsidies and firm financial performance: evidence from renewable energy manufacturing in China. Renew Energy 63(1):330–336
28. Zhang J, Marquis C, Qiao K (2016) Do political connections buffer firms from or bind firms to the government? A study of corporate charitable donations of Chinese firms. Organ Sci 27:1307–1324
29. Zhang Y (2015) The contingent value of social resources: entrepreneurs' use of debt-financing sources in western China. J Bus Ventur 30(3):390–406
30. Zhou W (2013) Political connections and entrepreneurial investment: evidence from China's transition economy. J Bus Ventur 28(2):299–315

Scenario-Based Location Arc Routing Problems: Introducing Mathematical Models

Alireza Amini[1], Reza Tavakkoli-Moghaddam[2(✉)], and Sadoullah Ebrahimnejad[3]

[1] School of Industrial Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran
[2] School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran
tavakoli@ut.ac.ir
[3] Department of Industrial Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran

Abstract. A location arc routing problem (LARP) is an important issue that finds the best locations of depots and routing simultaneously. It deals with a routing problem, in which demands are on arcs instead of nodes. Additionally, parameters may not be deterministic in real problems. Thus, this paper addresses an uncertain LARP through developing a deterministic mathematical model regarding to the respective literature and employing two scenario-based approaches. The objectives of the model are to minimize the maximum regret and minimize the mean and deviation of the objective function value (OFV). A numerical example is generated and the results analyze the performance of scenario-based models.

Keywords: Location arc routing · Mathematical model · Scenario-based · Regret

1 Introduction

An important cost of the most of companies is their logistic costs. It can be reduced through appropriate planning and designing supply chains. In fact, distribution networks in supply chains are highly important [9]. The location problem aims at finding optimal locations for facilities or centers. In fact, one of the most important factors in the success of a production unit is determining appropriate locations of sites and also, on a smaller scale, production facilities. Therefore, finding optimal or near-optimal solutions is essential. Vehicles routing, on the other hand, is among the most challenging issues in supply chain management. These two issues affect the cost of networks and supply chains. Hence, the location-routing problem (LRP)is defined to determine the location and routing decisions in at the same time. It is a combination of both location and routing optimization problems, whose decisions are made at the same

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_41

time. In recent years, several LRPs were presented. Considering the LRP, better logistics could be designed [4].

In general, routing issues may be defined on a graph of number of nodes and arcs. The LRP deals with routing problems where demands are on nodes. One related issue which has been considered in recent years by some researchers is the routing problem when demands belong to arcs instead of nodes. Arc routing is a special type of routing problems. In this content, researchers have taking into account the conditions and restrictions in real applications, a variety of models, and methods for solving these problems. Ghiani and Laporte [5] addressed a problem and solved it based on the concept of the rural postman problem (RPP) as one the earliest studies on arc routing problems. There are many researches in this filed. For example, Pia and Filippi [8] addressed a capacitated arc routing problem (CARP) to consider a waste collection problem and Beullens et al. [1] employed some heuristics for a periodic CARP.

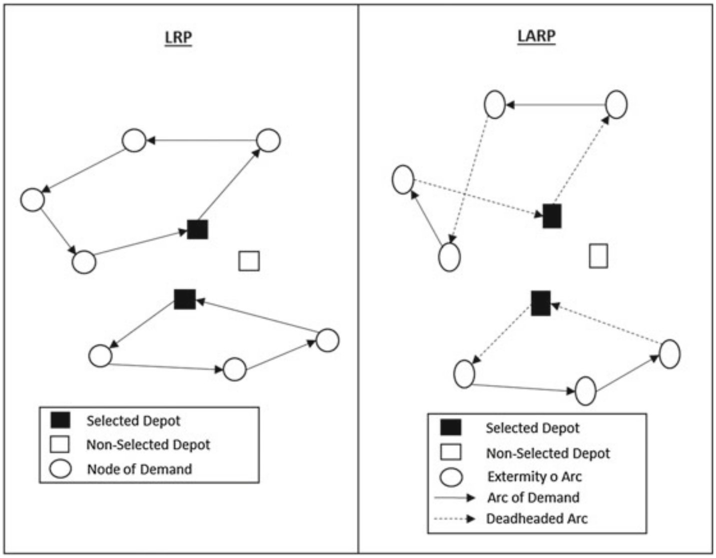


Fig. 1. LRP versus LARP

The location arc routing problem (LARP) is emerged when arc routing decisions are made along with location ones. The difference between the LRP and the LARP is displayed by Fig. 1. Waste collection, mail delivery, and telecommunication network design is among most well-known problems in which LARP could be addressed [7]. Also, some node routing problems could be considered as an arc routing one. For instance, when there are supermarkets as the nodes of demand, one can convert a node routing to an arc routing problem; because the street with N supermarkets is an arc with an aggregated demand from them. It might reduce the complexity of the problem. It means that a LARP could be defined for a lot of real applications. Firstly, Levy and Bodin [2] considered a LARP by which a mail delivery problem is taken into account. Then, Ghiani and Laporte [6] employed some heuristics to deal with a LARP. Lopes et al. [11] developed a mathematical model for LARP. In addition, Hashemi and Seifi [3] developed two mathematical models for single-depot and multiple-depot LARPs, respectively. One the latest works is what Riquelme-Rodrłguez et al. [10] presented which deals with a LARP with inventory constraints.

The paper on hand addresses a LARP by developing a mathematical model. Also, according to the uncertain nature of the data and information in the real world, we intend to build our research based on this fact. In our best of knowledge there is a gap in considering uncertainties in LARP. So, two scenario-based approaches are employed in the developed mathematical model. Section 2 defines the problem and notations and develops a deterministic mathematical model for LARP. Section 3 deals with two scenario-based approaches and modifies the mathematical model for each. In Sect. 4 a numerical example is analyzed based on the mathematical models. Finally, we conclude the paper.

2 Deterministic LARP Model

There is a complete graph $G = (V, A)$ where V and A denotes the set of vertices and set of arcs, respectively. V is consisted of K as the set of potential depots and R as the set of extremities of arcs, $\mathrm{V} = \mathrm{R} \cup K$. Each arc denoted by (i, j) which means that the arc begins form vertex i and ends at j. There is a given number of homogenous vehicles for each potential depot in a set called P. Then each depot is able to have $|P|$ separate tours. Some arcs have positive demands and the demand of arc(i, j) equals the demand of arc(j, i). This means that both of them are not required to be serviced and it is sufficient to meet the demand of one of them. The problem is finding optimal depots among the potential ones, assign vehicles and arcs to them, and traverse those arcs in an optimum routing. It is assumed that arcs with positive demands must be serviced once by a vehicle and also split delivery of demands is not allowed. Number of deadheaded traversing in each arc is not limited. In addition, tours must begin from a depot and end to the same one. In the deterministic model are parameters are given in advance. Following notations are introduced to develop the mathematical model.

Parameters:

F_k: Establishment cost of depot k;
C_{ij}^{k}: Servicing cost of required arc (i, j) by depot k;
H_{ij}^{k}: Deadheading cost of required arc (i, j) by depot k;
W: Hiring cost of each vehicle;
D_{ij}: Demand of arc$(i, j) \in R$;
Q: Vehicle's capacity;
B_k: Capacity of depot k.

Variables:

x_{ij}^{pk}: Binary variable which equals 1 if required arc(i, j) is served by vehicle p of depot k;
y_{ij}^{pk}: Number of deadheaded traverses in arc(i, j);
f_{ij}^{pk}: Departed flow form arc(i, j);
v_{ij}: Binary variable which equals 1 if arc(i, j) has a positive demand;
q_{pk}: Binary variable which equals 1 if vehicle p of depot k is established;
g_{pk}: Departed flow form depot k by vehicle p;
z_k: Binary variable which equals 1 if depot k is established.

Regarding the assumptions and notations the mathematical model for the deterministic LARP is developed as below. As a matter of fact, this model is inspired by what is developed in [3].

$$\min \theta = \sum_{k} z_k F_k + \sum_{i,j,p,k} x_{ij}^{pk} C_{ij}^{k} + \sum_{i,j,p,k} y_{ij}^{pk} H_{ij}^{k} + \sum_{p,k} q_{pk} W \tag{1}$$

$$\text{s.t.} \sum_{k \in K, p \in P} (x_{ij}^{pk} + x_{ji}^{pk}) = v_{ij}, \forall (i, j) \in R \tag{2}$$

$$\sum_{k \in K, p \in P} (x_{ij}^{pk} + x_{ji}^{pk}) = 0, \forall (i, j) \in K \tag{3}$$

$$\sum_{i \in V} (x_{ij}^{pk} + y_{ij}^{pk}) - \sum_{i \in V} (x_{ji}^{pk} + y_{ji}^{pk}) = 0, \forall j \in V, \forall p \in P, \forall k \in K \tag{4}$$

$$q_{pk} \leq z_k, \forall p \in P, \forall k \in K \tag{5}$$

$$\sum_{j \in R} (x_{kj}^{pk} + y_{kj}^{pk}) = q_{pk}, \forall p \in P, \forall k \in K \tag{6}$$

$$\sum_{j \in R} (x_{jk}^{pk} + y_{jk}^{pk}) = q_{pk}, \forall p \in P, \forall k \in K \tag{7}$$

$$x_{ij}^{pk} + y_{ij}^{pk} \leq M q_{pk}, \forall (i, j) \in V, \forall p \in P, \forall k \in K \tag{8}$$

$$x_{ii}^{pk} + y_{ii}^{pk} = 0, \forall i \in V, \forall p \in P, \forall k \in K \tag{9}$$

$$\sum_{i \in K, i \neq k} y_{ij}^{pk} = 0, \forall p \in P, \forall k \in K \tag{10}$$

$$x_{ij}^{pk} \leq D_{ij}, \forall (i, j) \in R, \forall p \in P, \forall k \in K \tag{11}$$

$$D_{ij} \leq M v_{ij}, \forall (i, j) \in R. \tag{12}$$

$$\sum_{p \in P} q_{pk} \leq |P|, \forall k \in K \tag{13}$$

$$g_{pk} \geq \sum_{i,j} D_{ij} x_{ij}^{pk} - M(1 - q_{pk}), \forall p \in P, \forall k \in K \tag{14}$$

$$g_{pk} \leq \sum_{i,j} D_{ij} x_{ij}^{pk} + M(1 - q_{pk}), \forall p \in P, \forall k \in K \tag{15}$$

$$g_{pk} \leq Q q_{pk}, \forall p \in P, \forall k \in K \tag{16}$$

$$\sum_{p \in P} g_{pk} \leq B_k, \forall p \in P, \forall k \in K \tag{17}$$

$$\sum_{i} f_{ij}^{pk} - \sum_{i} f_{ji}^{pk} = \sum_{i} D_{ji} x_{ji}^{pk}, \forall (i,j) \in V, \forall p \in P, \forall k \in K \tag{18}$$

$$f_{ij}^{pk} \leq Q(x_{ij}^{pk} + y_{ij}^{pk}), \forall (i,j) \in V, \forall p \in P, \forall k \in K \tag{19}$$

$$\sum_{j} f_{kj}^{pk} = g_{pk}, \forall j \in V, \forall p \in P, \forall k \in K \tag{20}$$

$$\sum_{j} f_{ik}^{pk} = 0, \forall i \in V, \forall p \in P, \forall k \in K \tag{21}$$

$$x_{ij}^{pk}, v_{ij}, z_k, q_{pk} \in \{0,1\}, \forall (i,j) \in V, \forall p \in P, \forall k \in K \tag{22}$$

$$y_{ij}^{pk} \geq 0 \& Integer, \forall (i,j) \in V, \forall p \in P, \forall k \in K \tag{23}$$

$$f_{ij}^{pk}, g_{pk} \geq 0, \forall (i,j) \in V, \forall p \in P, \forall k \in K. \tag{24}$$

Objective function (1) minimizes the total costs of opening depots, traversing, and hiring vehicles. Constraint (2) guarantees that required arc(i,j) must be served by one vehicle form one depot in one direction, i.e. from i to j or j to i. Constraint (3) ensures that traveling between depots is not allowed. Constraint (4) represents the continuity of tours, and Constraint (5) is in charge of vehicle assignments to depots could be considered if that depot is opened. Constraint (6) indicates that one arc must leave the depot to one arc of its tour at the beginning of the tour and, on the other hand, Constraint (7) shows that arc must enter the depot form one arc of its tour at the end of the tour. In Constraint (8) it is mentioned that x_{ij}^{pk} and y_{ij}^{pk} can be equal to 1 if they are assigned to the vehicle p of depot k, where M is a big sufficient positive number.

Constraint (9) guarantees that there is not any arc between one vertex and itself. Constraint (10) declares that tours are not allowed to end at a different depot from which they begin. Also, Constraints (11) and (12) find the arcs with positive demands. Regarding Constraint (13), maximum number of assigned vehicles to each depot cannot exceed the total number of vehicles. Constraints (14) and (15) measure the total amount of demands for each vehicle of each depot if it is opened. Constraints (16) and (17) regard the capacity limitations for opened depots and selected vehicles. Constraint (18) guarantees that the existing flows from arc(i,j) equal the entering ones to that arc minus its demand if there is one servicing traveling for that arc in that tour. Also, Constraint (19) shows that flows of arc(i, j) are zero if there is not any traversing. Although, flows are not required to be determined, Constraints (18) to (21) ensure not to have sub-tours. Finally, Constraints (20) and (21): existing flow form each opened depot by each vehicle equals the total amount of demands of that tour and entering flow to that depot is zero.

3 Scenario-Based LARP

Regarding the real problems in the world, data might not be exactly given. So, many problems have uncertain parameters. The paper on hand aims at employing two scenario-based approaches to consider uncertainties: (i) Minimization of maximum regret, (ii) Minimization of mean and deviation of objective function value. The scenario-based LARP is called SLARP in this paper. So the first Approach is named SlARP^{R} and the second one is named SlARP^{MD}. Also, let s denotes the index of scenarios form its set S from now on.

3.1 SlARP^{R} Mathematical Model

This approach begins with employing the LARP model to solve the problem under each scenario separately. Then an integrated mathematical model is required to obtain a unique solution which meets all requirements of all scenarios. Developing the integrated mathematical model, following variables are introduced.

$r^{\max}$: Maximum regret;
f_{ijs}^{pk}: Departed flow form arc (i, j) under scenario s;
g_{pks}: Departed flow form depot k by vehicle p scenario s.

In fact, despite of other variables, flows are dependent on scenarios. Let θ_s^* indicates the optimum value for objective function under scenario s $\in S$. In addition θ_s is the obtained OFV for scenario s $\in S$ when the integrated model is solved, Eq. (25). Then, the regret of scenario sS is the difference between θ_s^* and θ_s. $\theta_s \geq \theta_s^*$ for each scenario s $\in S$.

$$\theta_s = \left(\sum_k z_k F_{ks} + \sum_{i,j,p,k} x_{ij}^{pk} C_{ijs}^k + \sum_{i,j,p,k} y_{ij}^{pk} H_{ijs}^k + \sum_{p,k} q_{pk} W_s \right), \forall s \in S. \quad (25)$$

Considering LARP model, Eq. (25), and new variables, following model is developed as the integrated model for SLARP^{R}.

$$\min \ \theta^R = r^{\max}, \quad (26)$$

$$\text{s.t. } r^{\max} \geq \theta_s - \theta_s^*, \forall s \in S \quad (27)$$

$$x_{ij}^{pk} \leq D_{ijs}, \forall (i,j) \in R, \forall p \in P, \forall k \in K, \forall s \in S \quad (28)$$

$$D_{ijs} \leq M v_{ij}, \forall (i,j) \in R, \forall s \in S \quad (29)$$

$$g_{pks} \geq \sum_{i,j} D_{ijs} x_{ij}^{pk} - M(1 - q_{pk}), \forall s \in S \tag{30}$$

$$g_{pks} \leq \sum_{i,j} D_{ijs} x_{ij}^{pk} + M(1 - q_{pk}), \forall s \in S \tag{31}$$

$$g_{pks} \leq B_{ks}, \forall p \in P, \forall k \in K, \forall s \in S \tag{32}$$

$$g_{pks} \leq B_{ks}, \forall p \in P, \forall k \in K, \forall s \in S \tag{33}$$

$$\sum_{i} f_{ijs}^{pk} - \sum_{i} f_{jis}^{pk} = \sum_{i} D_{ijs} x_{ij}^{pk}, \forall j \in V, \forall p \in P, \forall k \in K, \forall s \in S \tag{34}$$

$$f_{ijs}^{pk} \leq Q_s (x_{ij}^{pk} + y_{ij}^{pk}), \forall (i,j) \in V, \forall p \in P, \forall k \in K, \forall s \in S \tag{35}$$

$$\sum_{j} f_{ijs}^{pk} = g_{pks}, \forall (i,j) \in V, \forall p \in P, \forall k \in K, \forall s \in S \tag{36}$$

$$\sum_{i} f_{ijs}^{pk} = 0, \forall (i,j) \in V, \forall p \in P, \forall k \in K, \forall s \in S \tag{37}$$

$$f_{ijs}^{pk}, g_{pks} \geq 0, \forall (i,j) \in V, \forall p \in P, \forall k \in K, \forall s \in S. \tag{38}$$

Constraints (2–10), (13), (22), (23).

Objective function (26) minimizes the maximum regret among all scenarios measured by Constraint (27). Other constraints are similar to those presented for the LARP model. Some of them (i.e., Constraints (28) to (38)) are required to be considered for each scenario separately.

3.2 SlARP^{MD} Mathematical Model

This approach is similar with SlARP^R in beginning with solving problems under scenarios separately and takes all of them into account in an integrated model. Its main difference with SlARP^R is in its objective function. Also, SlARP^{MD} needs to know the probability of occurrence of each scenario, P_s. SlARP^{MD} aims at minimizing the mean of OFV, $\overline{\theta}$, and its deviation form, σ_θ^2, where $\overline{\theta} = \sum_s P_s \theta_s$ and $d_\theta = \sum_s P_s \left|\theta_s - \sum_{s'} P_{s'} \theta_{s'}\right|$. So, this approach minimizes $\overline{\theta} + d_\theta$ in its integrated model. It is worth mentioning that deviation could be differently important for the different decision makers. Therefore, the weight λ is introduced for it. Now the integrated model minimizes $\overline{\theta} + d_\theta$ which is $\sum_s P_s \theta_s + \lambda \sum_s P_s \left|\theta_s - \sum_{s'} P_{s'} \theta_{s'}\right|$. The deviation term results in having a non-linear objective function. As a result to non-negative variables are employed to linearize the objective function: e_{1s} and e_{2s}. Now the term $\left|\theta_s - \sum_{s'} P_{s'} \theta_{s'}\right|$ in objective function is replaced with $e_{1s} + e_{2s}$ when $\theta_s - \sum_{s'} P_{s'} \theta_{s'}$ equals to $e_{1s} - e_{2s}$. Following model is the integrated model for SlARP^{MD}.

$$\min \ \theta^{MD} = \sum_{\mathrm{s}} P_s \theta_s + \lambda \sum_{s} P_s (e_{1s} + e_{2s}), \tag{39}$$

$$\text{s. t. } \theta_s - \sum_{\mathrm{s}'} P_{s'} \theta_{s'} = e_{1s} - e_{2s}, \forall s \in S. \tag{40}$$

Constraints (2–10), (13), (22), (23), (27-38).

4 Numerical Example

A numerical example is considered to analyze developed models. This example consists of three potential depots, two vehicles for each, and six extremities for arcs. We assume that there is a link between each pair of extremities and between each potential depot and each extremity. Potential depots belong to set $\{1, 2, 3\}$ and extremities are in set $\{4, 5, 6, 7, 8, 9\}$. In this example arcs there are positive demands between extremities $(4, 5)$, $(4, 7)$, $(4, 9)$, $(5, 7)$, $(5, 8)$, $(6, 7)$, $(6, 8)$, $(6, 9)$, and $(8, 9)$. Arcs with positive demands must be served in one direction. For instance, one of arcs 4-5 and 5-4 must be chosen for meeting the demand between extremities 4 and 5. Five parameters of the model are generated with patterns $W \sim \text{Uniform}(100, 200)$, $F \sim \text{Uniform}(400, 800)$, $C_{ij}^{k} \sim \text{Uniform}(50, 100)$, $H_{ij}^{k} \sim \text{Uniform}(10, 50)$, and $D_{ij} \sim \text{Uniform}(100, 1000)$. In addition, parameters of capacities are generated regarding the demands in a way that infeasibility does not occur.

Also, five scenarios are defined. Here, scenario $S1$ has the lowest amount of demands among all scenarios. Then, we assume that this scenario has the lowest amount of capacities and costs. So, scenario $S5$ has the highest amount for each parameter. All parameters of a scenario are generated by above-mentioned patterns through dividing the ranges of each parameter. For example, range (100,200) is assumed for W. This range is divided to five ranges to generate the scenarios. Thus, ranges (100,120), (121,140), (141,160), (161,180), and (181,200) are the ranges of the uniform distribution function of W for scenarios $S1$ to $S5$, respectively. This process is employed for each parameter.

Table 1. Optimal depots and routings for each scenario

Scenarios	Depot	Vehicle	Sequence	Deadheaded
S1	3	1	3-7-6-9-8-3	3-7, 8-3
	3	2	3-5-7-4-5-8-6-4-9-3	3-5, 6-4, 9-3
S2	2	1	2-7-4-9-8-6-2	2-7, 6-2
	3	1	3-8-5-7-6-9-3	3-8, 9-3
	3	2	3-4-5-3	3-4, 5-3
S3	2	1	2-4-9-6-8-5-8-9-2	2-4, 5-8, 9-2
	2	2	2-6-7-4-5-7-2	2-6, 7-2
S4	1	1	1-8-5-7-4-9-6-1	1-8, 4-9, 6-1
	3	1	3-7-6-8-9-4-5-3	3-7, 5-3
S5	3	1	3-6-9-4-5-4-7-5-8-3	3-6, 5-4, 8-3
	3	2	3-9-8-6-7-3	3-9, 7-3

The mathematical models are solved by GAMS optimization software. Table 1 presents the optimum locations of depots and routings for each scenario, separately. Also, Table 2 presents the optimum objective function value (OFV)

Table 2. Values of objective functions

	λ	P^{set}	θ_s					θ^R	θ^{MD}		
			$S1$	$S2$	$S3$	$S4$	$S5$		$\lambda=0$	$\lambda=0.5$	$\lambda=1$
$S1$	-	-	1168.00	-	-	-	-	-	-	-	-
$S2$	-	-	-	2137.00	-	-	-	-	-	-	-
$S3$	-	-	-	-	1757.00	-	-	-	-	-	-
$S4$	-	-	-	-	-	2545.00	-	-	-	-	-
$S5$	-	-	-	-	-	-	2244.00	-	-	-	-
θ^R	-	-	1811.00	2201.00	2533.00	2802.00	3263.00	1019.00	2522.00	3554.00	4586.00
θ^{MD}	0	$\alpha1$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2498.40	2713.56	2928.72
θ^{MD}	0	$\alpha2$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2243.70	2455.32	2666.94
θ^{MD}	0	$\alpha3$	1750.00	2161.00	2531.00	2763.00	3296.00	1052.00	2326.50	2508.00	2689.50
θ^{MD}	0	$\alpha4$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2497.40	2637.62	2777.84
θ^{MD}	0	$\alpha5$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2656.90	2823.56	2990.22
θ^{MD}	0	$\alpha6$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2765.90	2978.74	3191.58
θ^{MD}	0.5	$\alpha1$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2498.40	2713.56	2928.72
θ^{MD}	0.5	$\alpha2$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2243.70	2455.32	2666.94
θ^{MD}	0.5	$\alpha3$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2327.10	2505.36	2683.62
θ^{MD}	0.5	$\alpha4$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2497.40	2637.62	2777.84
θ^{MD}	0.5	$\alpha5$	1757.00	2173.00	2531.00	2747.00	3297.00	1053.00	2657.40	2821.16	2984.92
θ^{MD}	0.5	$\alpha6$	1811.00	2201.00	2533.00	2802.00	3263.00	1019.00	2773.40	2974.96	3176.52
θ^{MD}	1	$\alpha1$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2498.40	2713.56	2928.72
θ^{MD}	1	$\alpha2$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2243.70	2455.32	2666.94
θ^{MD}	1	$\alpha3$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2327.10	2505.36	2683.62
θ^{MD}	1	$\alpha4$	1752.00	2169.00	2520.00	2753.00	3298.00	1054.00	2497.40	2637.62	2777.84
θ^{MD}	1	$\alpha5$	1757.00	2173.00	2531.00	2747.00	3297.00	1054.00	2657.40	2821.16	2984.92
θ^{MD}	1	$\alpha6$	1740.00	2188.00	2542.00	2759.00	3290.00	1046.00	2769.00	2977.40	3185.80

of each scenario. We can find that scenarios do not have same results to each other. It means that the value of parameters affects the optimum decisions even when the arc with positive demands are same. Also it is found that if traversing costs of each tour are dependent to its depot, decision makers cannot simply declare that having multiple tours with one depot is better than having multiple depots. For example, although scenarios $S1$, $S3$, and $S5$ prefer to use two tours for the opened depot and not to open an extra one, $S4$ opens two depots with one tour for each while it can assign another vehicle to each of them and remove the other one. In fact, parameters of the problem indicate the best solution.

SlARP^R and SlARP^{MD} models have been taken into account after obtaining optimal solutions of scenarios. Regarding the objective function of SlARP^{MD}, a value for λ is required. This example assumes three different levels for λ: 0, 0.5, and 1. Also, this model needs to know the probability of each scenario in advance. Considering different conditions, six sets, called P^{set}, are given for the following probabilities of scenarios: $\alpha1 = \{0.2, 0.2, 0.2, 0.2, 0.2\}, \alpha2 = \{0.4, 0.2, 0.2, 0.1, 0.1\}, \alpha3 = \{0.2, 0.4, 0.2, 0.1, 0.1\}$, $\alpha4 = \{0.1, 0.2, 0.4, 0.2, 0.1\}, \alpha5 = \{0.1, 0.1, 0.2, 0.4, 0.2\}$, $\alpha6 = \{0.1, 0.1, 0.2, 0.2, 0.4\}$, where each set is defined as $\{\mathrm{P}_{s1},\mathrm{P}_{s2},\mathrm{P}_{s3},\mathrm{P}_{s4},\mathrm{P}_{s5}\}$. Table 2 indicates the OFV of models under different values for λ and set of probabilities and all equivalent OFVs for scenarios and other models. For instance, the optimum value of θ^R are fixed for scenarios and SlARP^{MD}

Table 3. Optimal depots and routings for the SlARP^R and two cases of the SlARP^{MD}

Model	Depot	Vehicle	Sequence	Deadheaded
SLARP^R	1	1	1-4-9-6-7-5-1	1-4, 5-1
	2	1	2-6-8-9-2	2-6, 9-2
	2	2	2-7-4-5-8-2	2-7, 8-2
SLARP$^{MD}(\lambda = 0, \alpha 1)$	1	1	1-8-5-7-4-5-1	1-8, 5-1
	3	1	3-9-8-6-7-3	3-9, 7-3
	3	2	3-4-9-6-3	3-4, 6-3
SLARP$^{MD}(\lambda = 1, \alpha 6)$	1	1	1-6-9-8-5-4-1	1-6, 4-1
	3	1	3-9-4-7-5-3	3-9, 5-3
	3	2	3-8-6-7-3	3-8, 7-3

model their OFVs are measured. It is worth mentioning that we assume that scenarios have same probabilities when we want to fix the SlARP^R optimum value of decision variables in SlARP^{MD}.

It can be figured out that each OFV is in its lowest level when its respective model is optimized. Obviously, some other problems reached the same value. For example, θ^{R^*} and it equals to the equivalent value for θ^R when SlARP^{MD} model is optimized with $\lambda = 0.5$ and $\alpha 6$. Also, one can find some optimum decisions in Table 3. It shows the effect of parameters and models in optimum solutions. Minimizing the maximum regret of scenarios, depots 1 and 2 are opened while depots 1 and 3 are opened when SlARP^{MD} model is considered. In addition, λ and P^{set} affect the routing of arcs. Minimizing total costs, the models prefer not to deadhead arcs as much as possible. Most deadheading traverses relate to the linking arcs between depots and extremities.

5 Conclusion

This paper aimed at considering uncertainties in a location arc routing problem (LARP). First, a deterministic mathematical model was developed. Then, two scenario-based approaches were employed to deal with the uncertainties of parameters. The bi-objectives were to minimize the maximum regret (SLARP^R) and minimize the mean and deviation of the objective function value (OFV) (SLARP^{MD}). It was required to find the optimal solution of each scenario in both approaches. The SLARP^R measured the distance between the OFV of each scenario to its optimal value (i.e., the regret of each scenario), and minimized the maximum one. On the other hand, the SLARP^{MD} minimizes the mean of OFVs of scenarios and their deviations from their optimum OFVs. A numerical example was generated to demonstrate the performance of models. Further studies may take the other approaches of uncertainty into account in the LARP.

References

1. Beullens P, Muyldermans L, Cattrysse D et al (2003) A guided local search heuristic for the capacitated arc routing problem. Eur J Oper Res 147(3):629–643
2. Bodin L, Levy L (1989) The arc oriented location routing problem. INFOR Inf Syst Oper Res 27(1):74–94
3. Doulabi SHH, Seifi A (2013) Lower and upper bounds for location-arc routing problems with vehicle capacity constraints. Eur J Oper Res 224(1):189–208
4. Drexl M, Schneider M (2014) A Survey of the Standard Location-Routing Problem. Publications of Darmstadt Technical University Institute for Business Studies
5. Ghiani G, Laporte G (1999) Eulerian location problems. Networks 34(4):291–302
6. Ghiani G, Laporte G (2001) Location-arc routing problems. Opserach 38(2):151–159
7. Nagy G, Salhi S (2007) Location-routing: issues, models and methods. Eur J Oper Res 177(2):649–672
8. Pia AD, Filippi C (2006) A variable neighborhood descent algorithm for a real waste collection problem with mobile depots. Int Trans Oper Res 13(2):125–141
9. Prodhon C, Prins C (2014) A survey of recent research on location-routing problems. Eur J Oper Res 238(1):1–17
10. Riquelme-Rodríguez JP, Gamache M, Langevin A (2016) Location arc routing problem with inventory constraints. Comput Oper Res 76:84–94
11. Rui BL, Fea Plastria (2014) Location-arc routing problem: heuristic approaches and test instances. Comput Oper Res 43(3):309–317

The Impact of Industrial Structure Change on Economic Growth

Hongxia Liu(✉), Zongtang Xie, and Xin Song

School of Economics, Northwest Minzu University,
Lanzhou 730030, Gansu, People's Republic of China
shirleygnd@163.com

Abstract. Taking case of Hebei Province, by building the regression model, this paper estimates and measures the rationalization and high-level of industrial structure effect on economic growth. The results show that the rationalization of industrial structure and advanced industrial structure can bring positive effects on economic growth. The effect of industrial structure rationalization on economic growth is greater than advanced industrial structure. In addition, the relationship between the rationalization of industrial structure and economic growth is basically stable in Hebei, while the relationship between advanced industrial structure and economic growth is instability. Finally, this paper puts forward some suggestions to optimize the industrial structure to promote economic growth.

Keywords: Industrial structure change · Economic growth, Rationalization of industrial structure · Advanced industrial structure

1 Introduction

Industrial structure and regional economic growth are interrelated and promote each other, the upgrading of industrial structure will promote the regional economic growth, regional economic growth process accompany with industrial structure change, reasonable evolution of industrial structure has become important symbol of economic upgrading and modernization, which can promote the potential power of regional economic growth. Domestic and foreign scholars make a lot of research on the industrial structure and economic growth. In the late 1960s, Kuznets explored the process of economic growth in the United States from 1948 to 1966, believing that industrial structural change was an important driving force of economic growth. In the 1980s, Chenery et al. found that economic growth was unbalanced, the production factors would flow from the low-yield sector to the high-yield sector, and the effect of industrial structure can drive economic growth [1]. At the beginning of this century, Peneder argued that the differences in productivity and its changing trends among production sectors would lead to the flow of the production factors in various sectors, the

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_42

resulting structural dividend could cause economic growth [6]. Perez and Freeman indicated the inherent relationship between industrial structure and the whole economic and social development model and the level [2]. Jorgenson discussed the origin of US economic growth and its international comparison, and revealed the role of the evolution of industrial structure [5]. Gan et al. took use of stochastic frontier production function to research and find that industrial structure and institutional factors could directly affect the economic scale, but also through the indirect impact of resource allocation function, and then they acted on the efficiency of output to promote regional economic growth [3].

2 Index Selection

2.1 Measurement of Industrial Structure Rationalization

Rationalization of the structure refers to the coupling relationship of input-output in, reflecting the rational use of resources and the degree of coordination among the various industries. The Theil index was originated from the entropy concept in information theory, which was used by Theil to calculate the degree of income gap in 1967. It is also scientific to use this index to measure the rationalization of the structure. Referring to the relevant research of Chunhui Gan et al. [3], this paper adopts the Theil index, the formula as following:

$$TL = \sum_{i=1}^{n} \left(\frac{Y_i}{Y}\right) \ln\left(\frac{Y_i}{Y} \Big/ \frac{L_I}{L}\right) = \sum_{i=1}^{n} \left(\frac{Y_i}{Y}\right) \ln\left(\frac{Y_i}{L_I} \Big/ \frac{Y}{L}\right),$$

TL is the Theil index, Y is GDP, Y_i is output value of various industries, L is total employment in a region, L_i is number of employment in various Industry. TL is equal to 0, the industrial structure is reasonable, TL is not 0, the economic deviates from the equilibrium state, the industrial structure is unreasonable. The closer the index value is 0, the more reasonable the structure is.

2.2 Measurement of Advanced Industrial Structure

The sophistication is the dynamic process of industrial structure, evolving from low-level to high-level. In this paper, the ratio of third industry and the second industry (TS) is used to measure the high grade, which can not only reflect the service-oriented tendencies of the structure, but also clearly indicate the development direction of industrial structure. Thus, using TS to measure the advanced industrial structure is rationality and scientific. The calculation formula is:

$$TS = \frac{Y_3}{Y_2},$$

where TS is output ratio of the third and second industrial, Y_3 is output value of the third industrial, Y_2 is output value of the second industrial. When TS rises, it shows economic is the evolution of services, industrial structure is in the continuous upgrading.

2.3 Measurement of Economic Growth

This paper chooses the Gross Domestic Product (GDP) as a measure of economic growth. Gross Domestic Product (GDP) refers to the market value of the final products and services which a country or a region uses production factors to produce in a given period. It not only reflects a country's economic performance, but also can reflect national strength and wealth.

3 Model Construction

In order to further explore the impact of industrial Structure rationalization and advanced Industrial Structure on the economic growth of Hebei Province, this paper uses multiple regression model, the model is set to

$$\ln G_i = \beta_1 + \beta_2 \ln TL_i + \beta_3 \ln TS_i + u_i.$$

In the model, G is regional GDP, TL is the Theil index, referring to the rationalization of industrial structure, TS is output ratio of the third and second industrial, to measure advanced industrial structure. In order to eliminate heteroscedasticity, each variable is taken as a natural logarithm, and u is the random error term, β_1, β_2, β_3 is parameter.

4 Empirical Analysis of Industrial Structure Change on Economic Growth in Hebei Province

4.1 The Characteristics of Industrial Structure Change in Hebei Province

This paper selects the relevant data from 1978 to 2012 in Hebei, and applies Eviews7.0 software to process the relevant data. The calculated results of industrial structure rationalization and advanced industrial structure are shown in Table 8.

Table 1 shows: Theil index mean is 0.2324, the standard deviation is 0.0517, the maximum and minimum are 0.3508, 0.1466. In 28 years, degree of industrial structure rationalization kept steady. TS appears overall upward trend, its mean is 0.6596, standard deviation is 0.0709, the maximum and minimum are 0.8160,0.5001, showing volatility rise (Table 2).

Figure 1 shows that from 1985 to 2012, TL changes from initial higher than 0.2 to later less than 0.2, showing that industrial structure is gradual in the reasonable trend. At the same time, TS changes from the initial 0.4 and 0.6 to 0.6 and 0.8 later, industrial structure is a high-level forward in Hebei Province. In addition, it should be emphasized that the correlation of TL and TS is not strong (as shown in Table 1), and the change trend of them is not the same. Thus, the interaction of TL and TS need not to be considered when analyzing the impact industrial structure change on economic growth (Table 3).

Table 1. Indicators values of industrial structure rationalization and advanced measurement in Hebei

Year	TL	TS	Year	TL	TS
1985	0.2252	0.5001	1999	0.2275	0.6943
1986	0.2268	0.511	2000	0.2518	0.6777
1987	0.2317	0.5012	2001	0.2409	0.7069
1988	0.2679	0.6667	2002	0.2383	0.7383
1989	0.2831	0.671	2003	0.2455	0.7139
1990	0.2706	0.7249	2004	0.2156	0.6607
1991	0.3481	0.816	2005	0.2216	0.6337
1992	0.3398	0.7821	2006	0.2219	0.6375
1993	0.3508	0.6383	2007	0.1924	0.6388
1994	0.2551	0.648	2008	0.1969	0.6063
1995	0.2023	0.6769	2009	0.179	0.6773
1996	0.1858	0.6533	2010	0.1713	0.6653
1997	0.197	0.6502	2011	0.1667	0.6462
1998	0.2065	0.6626	2012	0.1466	0.6702

Table 2. Descriptive statistics

	N	Maximum	Minimum	Mean	Standard deviation
VAR0001	28	0.3508	0.1466	0.2324	0.0517
VAR0002	28	0.816	0.5001	0.6596	0.0709

Table 3. Pearson correlation test of TL and TS

	TL	TS
TL Pearson correlation	1	0.094
Sig.tow-tailed	-	0.693
N	28	28
TS Pearson correlation	0.094	1
Sig.tow-tailed	0.693	-
N	28	28

4.2 Regression Analysis Results

(1) Stationary Test of Time Series
In reality, most of the time series variables are non-stationary, which may lead to pseudo-regression and invalid result in the regression analysis [4,7,8]. Therefore, before testing the co-integration relationship between GDP and industrial structure in Hebei, this paper adopts unit root to test stationary of time series.

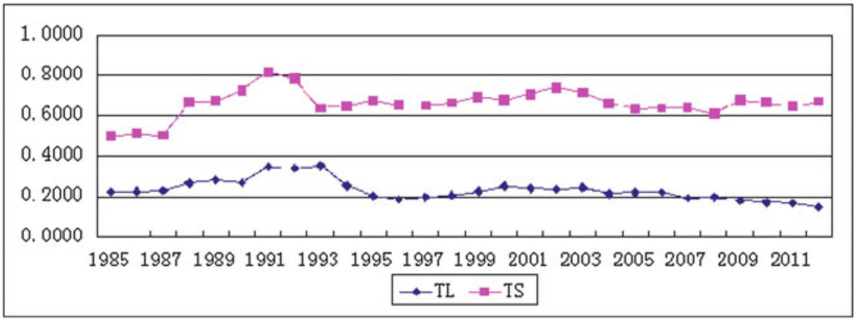


Fig. 1. Change trend of TL and TS

Table 4. Unit root test results

Variable	Stationary difference order	t-Statistic	Test critical values (10%)
lnGDP	level one	−3.316089	−3.233456
lnTL	level one	−4.016404	−3.233456
lnTS	level one	−4.341045	−3.254671

ADF test results show that the t-test values of lnGDP, $\ln TL$, $\ln TS$ are less than the critical value at the significance level of 10%, indicating that the null hypothesis is rejected at least confidence level of 90%, lnGDP and $\ln TL$, $\ln TS$ are integrated of order on. Thus, the long-term relationship between lnGDP and $\ln TL$, $\ln TS$ can be co-integration test.

(2) Co-integration Test
In the co-integration test, the Johansen maximum likelihood method is used to test the co-integration relationship among variables. For a number of non-stationary time series, if one linear combination is stationary, there exists a long-term equilibrium co-integration relationship among these variables. The trace statistic and the maximum eigenvalue statistic can be used to test. The test results are shown in Table 5:

Table 5. Co-integration test results (trace test)

Hypothesized trace = 0.05				
No. of CE(s)	Eigenvalue	Statistic	Critical value	Prob.**
None*	0.860236	68.96358	29.79707	0
At most 1*	0.363917	19.76856	15.49471	0.0107
At most 2*	0.28703	8.457905	3.841466	0.0036

Note: *indicates that the statistical significance is at 5% level, the same below. **indicates that the statistical significance is at 1% level.

Table 6. Co-integration test results

Hypothesized max-eigen = 0.05				
No. of CE(s)	Eigenvalue	Statistic	Critical Value	Prob.**
None*	0.860236	49.19502	21.13162	0
At most 1*	0.363917	11.31066	14.2646	0.0059
At most 2*	0.28703	8.457905	3.841466	0.0036

Note: *indicates that the statistical significance is at 5% level, the same below. **indicates that the statistical significance is at 1% level.

The results of co-integration test in Tables 5 and 6 show that the trace statistic and the largest eigenvalue statistic are larger than the critical value at the significance level of 5%, indicating that the hypothesis can be rejected at the 95% confidence level, indicating that there is co-integration relationship between the variables. Thus, there are more than two co-integration relationships among the three variables (lnGDP, $\ln TL$, $\ln TS$) at the significance level of 5%.

(3) Vector Error Correction Model

The above test shows that there exists co-integration relationship or long-term stability relationship among the variables. On this basis, the VEC model will be established to reflect the long-term static and short-term dynamic relationship. The results are as follows:

Table 7. Vector error correction model

Error correction	D(lnGDP)	D($\ln TL$)	D($\ln TS$)
CointEq1	0.009515	−0.000401	0.011106
	−0.00351	−0.00666	−0.0055
	[2.71014]	[−0.06021]	[2.01862]
D(lnGDP(−1))	0.231887	−0.763078	−0.239289
	−0.18074	−0.34268	−0.28324
	[1.28297]	[−2.22681]	[−0.84484]
D(lnTL(−1))	0.217572	0.041751	−0.072049
	−0.10565	−0.20031	−0.16557
	[2.05929]	[0.20843]	[−0.43517]
D(lnTS(−1))	−0.093157	0.507951	0.146124
	−0.14085	−0.26704	−0.22072
	[−0.66140]	[1.90214]	[0.66204]
C	0.117807	0.099722	0.046109
	−0.0301	−0.05707	−0.04717
	[3.91384]	[1.74742]	[0.97754]

Table 7 shows the estimated values of the parameters of the VEC model, where the value of CointEq1 represent the coefficient estimate of the error correction term. While the error equation can be derived from Table 4:

$$\text{Coint}Eq1 = D(\ln \text{GDP}) + 0.23 \ln \text{GDP}(-1) + 0.22 \ln TL(-1) - 0.09 \ln TS(-1) + 0.12.$$

This equation is deformed to the following equation:

$$D(\ln GDP) = \text{Coint}Eq1 - 0.23 \ln GDP(-1) - 0.22 \ln TL(-1) + 0.09 \ln TS(-1) - 0.12.$$

This error equation shows: when $\ln GDP$ changes each additional unit, the impact of $\ln GDP(-1)$ on it is -0.23 units, the impact of $\ln TL(-1)$ on it is -0.22 units, the impact of $\ln TS(-1)$ on it is 0.09 units. In addition, Theil index and GDP are negative correlation, $lnTL$ rises one unit, D($\ln GDP$) will drop 0.22 units; and TS and GDP show positive correlation, $\ln TS$ rises one unit, D($\ln GDP$) rises 0.09 units. This shows that the economy in Hebei Province will grow with the industrial structure rationalization and the advanced industrial structure, but its influence coefficient shows that the influence of industrial structure rationalization on the economy is higher than that of the advanced industry structure on economy in Hebei.

(4) Granger Causality Test

Table 8 shows: $\ln GDP$ and $\ln TL$ exist causal relationship, there is a mutually reinforcing relationship between industrial structure rationalization and regional economic growth; while advanced industrial structure and economic growth are only a single causal relationship, economic growth can promote advanced industrial structure, and advanced industrial structure can not promote economic growth. GDP in Heibei increased from 39.675 billion RMB in 1985 to 2.65751 trillion RMB in 2012, while the output value of the secondary industry and tertiary industry in Hebei Province has maintained a growing trend, indicating that Hebei's economic growth can promote industrial structure more reasonable. From 1985 to 2012, the second industry in Hebei still dominates in economy. But as time changes, the proportion of the first industry declines, the proportion of tertiary industry continues to rise, the industrial structure is more rationalized. This shows the stability of economic growth is mainly due to the rationalization

Table 8. Granger causality test results

Null hypothesis	Obs	F-Statistic	Prob
InTL is not the Granger cause of lnGDP	25	4.90145	0.0116
Ln GDP is not the Granger cause of lnTL		9.17695	0.0007
lnTS is not the Granger cause of lnGDP	25	0.95535	0.435
Ln GDP is not the Granger cause of lnTS		3.77428	0.0291
LnTS is not the Granger cause of lnTL	25	1.41804	0.2702
lnTL is not the Granger cause of lnTS		0.76488	0.5285

of the structure, but advanced industrial structure is only one aspect of economic growth effect, not to promote sustained economic growth.

5 Conclusion

This paper studies the impact of industrial structure dynamic change on economic growth from the angles of industrial structure rationalization and advanced industrial structure, and draws the conclusion that there is a long-term equilibrium trend between industrial structure change and economic growth, There is a stable relationship between the industrial structure rationalization and economic growth, while the instability of advanced industrial structure is more prominent; the industrial structure is the most important factor for the economic growth in Hebei. Evolution of industrial change in Hebei is in line with the general law, which will gradually move towards higher level.

Acknowledgements. This research was supported by the National Social Science Foundation of People's Republic of China (Grant NO. 11XMZ069), Scientific and Technological Project of Gansu Province of People's Republic of China (NO. 148ZCRA012) and the Special Fund of Fundamental Scientific Research Business Expense for Higher School of Central Government (Northwest Minzu University).

References

1. Chenery HB, Robinson S, Syrquin M et al (1986) Industrialization and growth: a comparative study, vol 4, pp 591–596
2. Freeman P (2000) Structural crises of adjustment, business cycles and investment behavior. Tech Change Econ Theor 38(6):38–66
3. Gan C, Zheng R, Yu D (2011) The impact of china's industrial structure change on economic wth and volatility. Struct Change Econ Dyn 5(4):4–18 (in Chinese)
4. Ghosh SK (2015) The effect of imperfect production in an economic production lotsize model. Int J Manag Sci Eng Manag 10(4):288–296
5. Jorgenson DW, Ho MS, Stiroh KJ (2003) Lessons from the us growth resurgence. J Policy Model 25(5):453–470
6. Peneder M (2003) Industrial structure and aggregate growth. Struct Change Econ Dyn 14(4):427–448
7. Tiemei G (2009) Economic analysis methodology and modeling - eviews application and example. Beijing Tsinghua Univ 1:126–165
8. Tripathi RP (2016) Economic ordering policies under credit financing for deteriorating items with stock-dependent demand rate using discounted cash flow approach. Int J Manag Sci Eng Manag 12:111–118

Disclosure Behavior of Annual Reports of Listed Companies Under Digital Worship

Xiaojing Xu(✉)

Chongqing Technology and Business University,
Chongqing 400067, The People's Republic of China
34144333@qq.com

Abstract. Digital worship is widely used in China. Does this phenomenon also affects the China's stock market? Does Digital Worship influence the choice of disclosure time of listed companies' annual reports? Through the empirical analysis for stock market of China, it founds that the listed companies' behavior of timing of annual reports released is affected by the short-term influence of digital worship, which is a new market phenomenon. The research of this paper has reference significance for investment decision-making, risk management and understanding the behavior of listed companies and investors.

Keywords: Digital worship · Annual report disclosure · Auspicious day

1 Introduction

Digital worship refers to the number of people to avoid or love [2]. The digital worship is widely used in China, such as the choice of telephone numbers, business opening date selection and so on. In this digital influence, people think that figure "8" seems to be "fatten" homophonic, meaning fortune. People think that the number "4" seems to be "dead" homophonic, and quite avoided. Such as the 2008 Beijing Olympic Games elected in August 8th at 8: 8 pm to open, and Alibaba Group was originally scheduled for August 8^{th}, 2014 to be on the United States stock market. Many studies have proved that there is a phenomenon of worshiping number 8 in China's stock market. In this paper the question posed is whether this digital worship exists in the timing of listed company's annual report disclosure? If so, does it have a significant impact on the market? Specifically, on the one hand, if listed companies are affected by the number of worship, whether they will choose date with the lucky number to publish annual report. On the other hand, if the investors' behavior is also affected by the digital worship, whether the listed companies will make good use of selecting the auspicious figures to disclose annual reports to get more investors' attention? If the digital worship has a impact on the disclosure behavior, it will be a new market vision.

The research of this paper has reference significance for investment decision-making, risk management and understanding the behavior of listed companies

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_43

and investors. The article will use the method of empirical analysis, with the date ending in number 8 and number 4 as the starting point to demonstrate whether the listed companies behavior of timing of annual report disclosure is affected by the digital worship. The remainder of the paper is structured as follows: The second part briefly reviews the literature. The third section describes the data and makes assumptions. The fourth part introduces the econometric model and empirical analysis. The fifth part summarizes the full text and makes recommendations.

2 Related Literature

Cao, Li et al. [2] studied China is a very particular about the "homophonic culture" of the country, 8 homonyms are "fatten", meaning fortune. And 4 homophonic is "dead", by a lot of people deliberately avoided. Xu [8] pointed out that the Chinese people believe that the numbers can predict good and bad. With social development, the number will have a new reference meaning and more distinct meaning, the number "8" is repeatedly subject to people's heat Holding. Shen [6] studied the symbolization and illusion of money, and discussed the influence of digital worship on modern wealth. In the stock market, Zhao and Wu [9] selected 569 non-financial stocks in Shanghai stock market as the object of study, and got code ending with 8 stock price high, its price-earnings ratio on IPO day and the following year higher than other mantissa stocks. Philip and Jason [1] found that the stock valuation of Shanghai and Shenzhen stock market there is a number of worship under the cluster effect, the price median with 8 appears twice the frequency of 4, because the Chinese inherent values that the number 8 is auspicious number, 4 is unlucky numbers. The above document proves that the stock code and the price contains the auspicious number will have an impact on the stock price. Liu, Li and Yang [4] found that auspicious number preference will significantly affect the relationship between the systematic risk of stock collapse and the expected rate of return. The above literature proves that there is a phenomenon of digital worship in many areas of Chinese stock market, and it has significant influence on the stock price and the expected return rate. However the number of worship will affect the information disclosure of listed companies and investors decision-making behavior, there is no relevant literature, which is the research direction of this paper.

Human decision-making process is not based on purely rational, but there will be a series of cognitive bias [5]. In a given context, people's decision-making behavior depends not only on the problem itself, but also on the way the problem is expressed. Han and Xu [3] found that listed companies significantly lower or early disclosure of the disclosure of annual reports will receive more attention to investors, which reveals the listed companies through the choice of time to disclose the annual report of an important channel for market value management. Zhou and Huang [10] found that the disclosure of our annual report on the stock market appeared "week preference" and "get together preferences". To disclose the good news on Tuesday, the stock excess return is higher. But the disclosure of bad news on Saturday, the stock is also higher than the excess rate of return. The stock has a positive impact on excess returns which the company chose the time when a number of annual reports published at the same time to disclose the bad news, in order to avoid the attention of investors. Wang and Wang [7] studied the impression management theory of information disclosure strategy of listed companies, and analyzed how listed companies can maximize their own benefits through management information disclosure. Thus, listed companies have strategy in the choice of disclosure of the timing, and they choose to publish are going to catch the chance of timing to attract and distract the attention of investors. This paper attempts to prove whether the listed companies use the digital worship effect and select the mantissa is the auspicious number of dates in order to achieve the effect of investor concern. In short, digital worship and disclosure strategy belong to the category of behavioral finance, and it is also a kind of market vision on the stock market.

3 Data and Assumptions

We collect and analyze the 2011–2015 Shanghai and Shenzhen A-share annual report disclosure date (a total of 12,610 samples) were statistics, the results shown in Fig. 1.

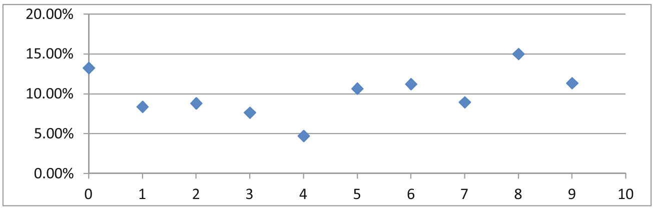


Fig. 1. 2011–2015 Annual report announcement date mantissa chart

As shown in the above figure, the number which listed companies choose to publish the annual report in the date of mantissa 8 is the most, accounting for 15.01%. But the date selected in the mantissa of 4 is the least, accounting for 4.72%. The results shows that the date listed companies choose to publish the annual report is Imbalance. There is a phenomenon of digital worship. However, how much dose this phenomenon impact the stock market? In particular, if the investor has the behavior of digital worship, the listed company chooses the auspicious date to disclose the annual report in order to get more attention from the investors when it chooses the disclosure time of the annual report. How to measure the annual report after the release of more investors are concerned about? This paper selects the 5-day yield index, which reflects the short-term price volatility of the stock market. If the five-day rate of return is greater than zero, then the stock received the attention of investors and investment, so that the stock prices up in the short term. Conversely, the opposite is true.

In order to facilitate the study, this paper chooses the date of the mantissa 8 as the auspicious date and the date of the mantissa 4 as the unlucky date to analyze whether the digital worship on the date will affect the disclosure time of the listed company. So this paper puts forward the hypothesis,

Hypothesis 1 (H1): If it choose the date of the mantissa 8 to publish annual report, the five-day stock returns will have a positive effect.

Hypothesis 2 (H2): If it choose the date of mantissa 4 to publish annual report, the five-day return has a reverse effect.

On the other hand, listed companies are also under the digital worship of praying auspicious, especially the annual report is bad news. They may choose to disclose the annual report in the auspicious date so that corporate earnings have a positive effect. The article sets that when the net profit margin is less than or equal to last year's net profit margin is bad news, when the net profit margin higher than last year's net profit margin is good news. Therefore,

Hypothesis 3 (H3): If the listed company which has good news in annual report select the date of mantissa 8 to disclose annual report, not only have a positive effect on the company's five-day yield, but also have higher five-day yield compared to other days.

Hypothesis 4 (H4): If the listed company which has good news in annual report select the date of mantissa 8 to disclose annual report, not only have a positive effect on the company's five-day yield, but also have higher five-day yield compared to other days.

4 Empirical Analysis

4.1 Data Selection and Data Sources

This data selection follows these principles: (1) As the annual reports always publish in the next year's January to April, so financial data selected from 2010 to 2014 annual report. And the yield selected from 2011 to 2015 annual report after the publication of five-day yield. Since the 2014 annual report is published in 2015, so the annual report data is from 2014, and the yield data is collect in 2015 after the release of annual report. (2) Removing some of the outliers and missing values; (3) The independent variables of the maximum and minimum of 3% flattening treatment. All the data in this article comes from the database of Guotai'an. The empirical analysis is based on Eviews6.0.

4.2 Variable Selection

(1) Dependent Variable

Five-day rate of return $(R_{it}) = \frac{P_1 - P_2}{P_2} \times 100\%$.

P_1 is the annual report after the fifth trading day closing price. P_2 is the annual report published on the day closing price (if the annual report is published after the end of the trading or in the non-trading day, we select the next trading day closing price.

(2) Independent variable
Assets-liability ratio (X_{1it}) = Total liabilities/Total assets × 100%.

Operating profit margin (X_{2it}) = Operating profit/Total business income × 100%.

Total assets growth rate (X_{3it}) = Total assets growth/Total assets at the beginning of the year × 100%.

Net cash flow per share from operating activities (X_{4it}) = Net cash flow from operating activities/Total share capital.

(3) Virtual Variable
This paper sets up the virtual variable to study whether the annual report can attract investors' attention and can cause the significant change of the 5-day rate of return. In order to study the annual report on the end of the mantissa of 8 or 4 on the impact of the five-day rate, we set as following:

Firstly, the annual report published on the end of 8 is 1, not the end of 8 is 0, with DV1 it said. Secondly, Announcement date to 4 at the end is 1, not the end of the 4 is 0, with DV2 it said. Thirdly if the listed company's net profit growth rate is less than or equal to last year is bad news, then the net profit growth rate is greater than last year is good news. The good news is 1, the other is 0, with DV3 it said. Finally, Annual bad news is published 1, the other is 0, with DV4 it said.

4.3 Build Models

This article refers to the requirements of the panel data model, the model is as follows:

Model 1 $R_{it} = X_{1it} + X_{2it} + X_{3it} + X_{4it} + DV_{1it} + \mu_{it}$,
Model 2 $R_{it} = X_{1it} + X_{2it} + X_{3it} + X_{4it} + DV_{1it}DV_{3it} + \mu_{it}$,
Model 3 $R_{it} = X_{1it} + X_{2it} + X_{3it} + X_{4it} + DV_{1it}DV_{4it} + \mu_{it}$,
Model 4 $R_{it} = X_{1it} + X_{2it} + X_{3it} + X_{4it} + DV_{2it} + \mu_{it}$,
Model 5 $R_{it} = X_{1it} + X_{2it} + X_{3it} + X_{4it} + DV_{2it}DV_{3it} + \mu_{it}$,
Model 6 $R_{it} = X_{1it} + X_{2it} + X_{3it} + X_{4it} + DV_{2it}DV_{4it} + \mu_{it}$.

R_{it} is the five-day rate of return, X_{1it} is on behalf of the asset-liability ratio, X_{2it} it represents operating profit margin, X_{3it} represents the total assets growth rate, X_{4it} represents the net cash flow per share of operating activities, $DV_{kit}(k = 1, 2, 3, 4)$ represents the dummy variables, μ_{it} represents the random error term.

4.4 Descriptive Statistics

During the inspection period, the five-day average yield is about 0.2297%, greater than zero. This shows that after the publication of the annual report, the majority of listed companies received the attention and recognition of the market, and stock prices rise. The maximum value is about 48.2540% and the minimum value is −50.2257%. This shows that the annual report after the release of the stock market response is very rapid, and investors can judge the tendency of markets (Table 1).

Table 1. Variable descriptive statistics

	R_{it}	X_{1it}	X_{2it}	X_{3it}	X_{4it}
Average value	0.2297	45.6519	8.564	16.8504	0.3557
Median	0	45.8901	7.064	10.5624	0.2712
The maximum	48.254	92.0426	42.0043	118.4737	2.4154
Minimum value	−50.2257	6.6931	−30.0346	−20.5382	−1.2789
The number of observation	8805	12508	12273	11521	12508
The number of cross sections	2433	2487	2451	2438	2487

Table 2. Empirical results

	Announcement date mantissa is 8 combinations			Announcement date mantissa is 4 combinations		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
X_{1it}	0.0108***	0.0089***	0.0088***	0.0111***	0.0081***	0.0073***
X_{2it}	−0.0077***	−0.0055***	−0.0054***	−0.0006	−0.0026	−0.0042*
X_{3it}	−0.0025**	−0.0040***	−0.0045***	−0.0048***	−0.0046***	−0.0054***
X_{4it}	0.0342*	0.0333***	0.0427***	−0.0192	0.0268	0.0126***
DV_{1it}	0.9320***	-	-	-	-	-
$DV_{1it}DV_{3it}$	-	1.0161***	-	-	-	-
$DV_{1it}DV_{4it}$	-	-	0.8512***	-	-	-
DV_{2it}	-	-	-	−1.6410***	-	-
$DV_{2it}DV_{3it}$	-	-	-	-	−1.5512***	-
$DV_{2it}DV_{4it}$	-	-	-	-	-	−1.8321***
C	−0.1687***	0.0798***	0.0876***	0.0026	0.1321***	0.2121***
R_2	22.01%	19.89%	88.11%	15.89%	2.43%	1.65%
$Fstatistic$	417.53***	351.01***	10201.19***	279.48***	31.98***	24.78***

Note: *, **, ***respectively means significant at 10%, 5%, 1% level

4.5 Empirical Analysis

In this paper, a hybrid panel data model is used. Using the cross-section weighting method, the empirical results are shown in Table 2.

4.6 Result Analysis

Based on the empirical results in Table 2, the analysis of the various research hypotheses is as follows:

Firstly, the DV1 coefficient of model 1 is 0.9320, which indicates that the listed company has positive effect on its 5-day rate of return on the date of the ending date of 8, and compared to other dates, which can improve the 5-day yield of 0.9320 units. So hypothesis 1 proves.

Secondly, as can be seen from model 2, the good news is better than other dates by 1.0161 units of the five-day rate of return on the date of tail number 8. So the good news announced in the date of tail number 8 can get more investors attention. So *Hypothesis 2* proves.

Thirdly, from Model 3 we can see that listed companies can increase the yield of 0.8512 units by 8 days, which means that bad news can weaken the negative effect of the stock market when the date is 8. So *Hypothesis 4* gives evidence. Finally, in the announcement date mantissa is 4 combinations, the model four DV2 coefficient is -1.6410. It indicated that the listed company in the date with tail number of 4 published annual report will reduce the 1.6410 unit rate of return. So *Hypothesis 2* gives evidence. From Model 5, we can see that the good news is released on the date ending on the 4^{th}, and the yield is reduced by 1.5512 units compared with other dates. From Model 6, we can see that the bad news is released on the date ending on the 4th, which will reduce the 1.8321 unit yield compared with other dates.

5 Conclusion and Suggestion

The influence of digital worship on the stock market belongs to the category of behavioral finance. It is also a kind of market vision and a hot topic in the field of finance research. From the perspective of digital worship, this paper analyzes the motivations of time choice of disclosure of listed companies' annual reports, and draws the following conclusions.

First of all, the annual disclosure of listed companies by the impact of China's digital culture, there is a certain tendency to auspicious. The number of listed companies which select the date with mantissa of 4 to publish annual report is the lowest. But the number of listed companies which select date with mantissa of 8 is the highest.

Second, whether the annual report is bad news or good news of listed companies, choosing the mantissa of the auspicious date to publish the annual report has a positive effect on five-day yield, and the positive effect of the good news is more obvious. Whether the annual report is good news or bad news of the listed companies, choosing the date mantissa of 4 to publish annual report has a reverse effect on five-day rate of return. This fully shows that digital worship behavior exists in the listed companies when they choose the time to announce annual report. While listed companies which has the bad news to use digital worship in the market, and choose a auspicious day to publish annual report in order to dilute the fundamentals of company unfavorable factors. On the other hand, investor behavior is also affected by the number of worship.

Last but not the least, we give the suggestions. The paper chooses the 5-day yield index as an empirical analysis, which is a short-term indicator. Therefore, listed companies in order to continue to receive the attention and recognition of investors should improve the quality of annual reports. Investors should also be aware of the impact on trading behavior which comes from the cultural and psychological factors, and finally make more rational investment decisions.

Acknowledgements. This paper is supported by The Ministry of Education of China and Doctoral Program of Specialized Research Project Fund Project (No. 20120184110021).

References

1. Brown P, Mitchell J (2004) Culture and stock price clustering: evidence from the peoples' republic of China. Soc Sci Electron Publ 16(1C2):95–120
2. Cao B (2015) Developmental characters and reason analysis of number preference. Adv Psychol 05(11):640–647
3. Han G, Xu H (2016) Operating performance of listed companies, disclosure of annual report and investors' concerns. China's Econ Probl 4:17–34
4. Li L (2015) Cognitive bias of compound proportion relation and its interpretation verification. PhD thesis, East China Normal University, Shanghai
5. Shen G (2016) Financialization, symbolism and illusion of money. J Wuhan Univ Sci Technol (Soc Sci Ed) 18:478–482
6. Shengyao Liu LY, Yunhong Y (2016) Collapse of Chinese stock market systematic risk and investor behavior preference. Finan Res 2:55–70
7. Wang X, Wang Y (2006) Theoretical basis of information disclosure strategy of listed companies. Auditing Econ Res 2:84–87
8. Xu L (2012) A comparison of connotations of Chinese and western digital cultures. J Zhejiang Univ Sci Technol 24:115–127
9. Zhao WF (2009) Asset price vision under digital worship. Econ Res 6:129–141
10. Zhou J, Huang D (2011) Investors' limited attention and time choice of listed companies' announcements. Secur Mark Guide 5:53–60

Energy Finance in Promoting Renewable Energy Sustainable Development in China: The Channels and Means of Financing

Lei Xu[1(✉)], Yanfei Deng[2], Yuan Yuan[3], and Karen Mancl[4]

[1] School of Economics, Xihua University, Chengdu 610039, People's Republic of China
leihsu@163.com

[2] School of Management, Southwest University for Nationalities, Chengdu 610064, People's Republic of China

[3] China Power Construction Corporation, Sichuan Electric Power Design and Consulting Co., Ltd., Chengdu, People's Republic of China

[4] College of Engineering, Ohio State University, Columbus, USA

Abstract. As global climate warming and the continued use of fossil energy are unsustainable, it has become imperative to accelerate the development of renewable energy. As one of the largest developing countries, this is especially true for future sustainable development in China; therefore, focusing on renewable energy can significantly reduce its dependence on imported fossil fuels. With policy support, China's renewable energy industries have had rapid development and are now at the international forefront level in some fields. However, the renewable energy industries are facing challenges, particularly in terms of energy finance. Basically, the renewable energy industry in China is facing a contradictory dilemma because of funding shortages and inefficient investment strategies, due in part to government-centered renewable energy investment and financing. In view of this, the renewable energy financing and the corresponding polices in China are discussed in detail in this paper. Issues of energy finance are also analyzed.

Keywords: Energy finance · Renewable energy · Financing channels · Financing means · Sustainable development

1 Introduction

Renewable energy (RE) has become the fundamental direction and core content of the global energy transformation. As renewable energy is environmentally friendly, low-carbon energy is the main force behind the technological revolution in energy and has become a major emerging industry growth point. China enacted a series of polices to promote sustainable RE development to increase

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_44

investment in RE industries and to gradually eliminate the obstacles to RE sustainable development. These government actions have improved the RE subsidy mechanism and incentives to ensure that renewable energy generation maintains a steady upward development trend. As China is only in the middle stages of industrialization and urbanization, energy demands continue to grow. The average annual RE installed capacity is expected to reach 2600 kilowatts or more every year in the future. There is a need, for example, for new investment in the solar RE industry of around 190 billion CNY to meet the goals of annual new installed capacity of 20 million kilowatts. Wind and solar power development together needs future annual investments of at least 400 billion CNY, with a total RE investment of 2 trillion CNY required for the "13*th* Five-Year" period [10].

International experts have now categorized RE into traditional and new, with the former referring to giant hydropower and directly burnt biomass and the latter referring to small hydropower, solar energy, wind energy, biomass energy, geothermal energy and others [3]. The renewable energies discussed in this article fall into the latter category. There are abundant RE resources in China; therefore, to make good use of these RE resources, China needs to promote the development of large new energy resources. Seeking to increase RE use and deployment to encourage a move towards sustainable development, governments have developed and implemented RE support policies. However, as the renewable energy sector grows, it is facing increasingly more complex investment, financing and policy challenges. Because of the initial high costs of renewable energy projects, market development requires focused policy support. If RE financing mechanisms are incomplete or poorly focused, renewable energy project financing has higher costs and attracts less investment; therefore, innovative investment and financing mechanisms are needed to break through the RE financing bottleneck and the continued reliance on government guidance and incentives and to encourage financial market innovations and policy incentives. Existing studies have found that renewable energy laws and regulations as well as incentive policies are an important driving force in the rapid RE development in China [1,5] such as incentive policies, financial subsidy policies, tax rebates and exemption policies, preferential feed-in tariff policies and technical support policies. China's current renewable energy policies include assistance with development plans, industrial guidance, technical support, legal responsibility and popularization [4,11]. Based on a questionnaire, Lam et al. found that the most important drivers of RE development in China were government financial assistance, technical support, feed-in-tariff policies and tax incentives.

2 The Driving Forces of RE Development

Strategically, it is important for China to develop domestic renewable energy manufacturing and promote RE applications for several reasons. First, the promotion and development of RE can ensure that China achieves its goal of 15% non-fossil energy by 2020 and 20% by 2030. Second, RE gives developing countries an important opportunity to catch up with developed countries in terms of

RE production capacity and to improve their energy structures [9]. Therefore, with these goals in mind, it is important to assess China's renewable energy financing and policies issues to understand how such large scale investments in renewable energies have been achieved, the major renewable power generation investment and financing channels, the types and effectiveness of Chinese government financing incentive mechanisms, the characteristics of China's renewable energy and whether the current policy mechanisms are suitable for the future development of renewable energy. These issues are fully discussed in this paper.

Seeking to satisfy the growing energy needs and to reduce greenhouse gas emissions, it is important that more RE resources be harnessed. However, the wider consumption of RE resources is closely related to financial support. Government financial support mechanisms and appropriate finance policies are important to enhance RE sustainable development. It has been widely recognized that renewable energy policies, such as incentive policies financial subsidy policies, tax rebates and exemption policies, preferential feed-in tariff policies, technical support policies, development plans, industrial guidance and technical support, price incentives, legal responsibility and popularization, are essential for the promotion of renewable energy development. The renewable energy market share has been expanding and in recent years, China's renewable energy has had significant development because of the implementation of the Renewable Energy Law and related policies. Existing studies have shown that incentive policies have been an important driving force behind the rapid RE development in China [8].

3 RE Financing Channels in China

The Energy foundation and the National Development and Reform Commission predicted that from 2005 to 2020, China's needed to invest 18 trillion CNY into the energy sector, with new energy, energy saving and environmental protection measures needing about 7 trillion CNY [6]. Therefore, government subsidies and investment alone is clearly insufficient, so full use needs to be made of market financing using financial channels to resolve the funding shortages. It has been pointed out that the development of appropriate financing channels and instruments for both end users and the industry was one of the drivers for increased investment [2]. This issue is analyzed in this part of the paper.

3.1 Bank Financing

Bank financing, such as policy-related loans and commercial finance, is an important channel for the renewable energy industry. To implement the country's new energy industry policy, the National Development Bank strengthened financial support for the new energy industry. Currently, bank financing plays a significant role in facilitating investment into the renewable energy sector. The World Bank, the Asian Development Bank and other international financial organizations provide low interest concessional loans to support investment into Chinese energy sources, transportation, environmental protection, public utilities and other non-profit projects. Of these, energy saving and emissions reduction have attracted the most finance. The International Bank for Reconstruction and Development also provides market-based (hard) loans and guarantees for Chinese renewable energy projects to increase the share of clean energy.

3.2 International Financing

International financing sources have played a major role as incubator funds for the development of rural renewable energy in China, with many renewable energy projects in China having been financed by multilateral and bilateral organizations. In China the internationally funding projects have included: "Capacity building for the rapid commercialization of renewable energy", a "China renewable energy development project", a "solar village project", a "wind power project in China's Hubei" and a "clean energy research project". International financing channels for China's renewable energy sector include international financial institution loans, inter-government loans, non-government loans and foreign investment, overseas Chinese investment in the territory, the Clean Development Mechanism (CDM) and the Carbon Abatement Fund.

3.3 Public-Private Partnerships

Public-private partnerships involve cooperation between a public sector authority and the private sector for the financing of revitalization, management or infrastructure maintenance projects including renewable energy projects. Private and public capital partnerships are generally joint ventures in which the entrepreneurs and the government work together to provide a faster and more efficient implementation of infrastructure projects. Also, because of the private capital injection, the assets that the public sector would otherwise invest in infrastructure projects in the traditional way can now be redirected towards creating social wealth. Additionally, there is a growing awareness that cooperation with the private sector in such public-private partnership projects has a number of advantages: (1) acceleration, (2) more rapid implementation, (3) reduced whole life infrastructure provision costs, (4) better risk allocation, (5) better incentives to perform, and (6) improved service quality.

3.4 Government Financing

Government financing also plays an important role in financing renewable energy in China, with the government taking a leading position in renewable energy sector financing. Government finance has been the most important channel for the financing of renewable energy projects through the implementation of national programs and the supply of grants to subsidize the capital costs of equipment, loans, loan guarantees and tax incentives. The Chinese government has a played a significant leading role in financing rural renewable energy through national initiatives and the provision of financial incentives. The Chinese Township Electrification Program was the first nationally based implementation of renewable technologies to supply electricity to rural populations and is one of the largest renewable energy-based rural electrification programs in the world. The development of such projects demonstrates the robust and sustainable renewable energy infrastructure in China, especially for solar energy.

3.5 Capital-Market Financing

Capital-market financing is a channel for new energy enterprises to raise funds as it can effectively solve fund shortage problems in a relatively short period of time. However, China's multi-level capital market is still somewhat behind as the capital market, stock market, bond market and investment products are not yet rich enough. China's Shanghai and Shenzhen stock exchange markets are very similar in terms of their organization, listing standards, trading patterns and regulatory bodies. The most direct way to utilize bonds for RE financing would be for RE companies to issue corporate bonds. China has had several successful case-studies for raising investment funds for the RE sector. Green bonds have emerged as a potential financing channel for RE, in which the issuer commits to the use of bond proceeds only for environmentally-friendly products. The key differentiating feature is that a green bond is marketed as green to investors and their proceeds are restricted to green projects. Issuing a green bond also increases the visibility of the bond to conventional investors. The ability to highlight the green attributes of green bonds could also improve overall publicity and improve an organization's image, thereby further broadening their access to capital.

4 RE Financing Means in China

Financing means, such as corporate financing, project financing, Third party financing and financing leases are relatively backward in China compared to developed countries [7].

4.1 Project Financing

Project financing has become one of the most important financing means for Chinese state-owned enterprises. A common project financing method is the Build-operate-transfer (BOT) method, with concessions having been granted to attract foreign and private investment. BOT was introduced as an alternative project financing model to motivate private investment participation. The government and power supplier sign a power purchase agreement, and to ensure that the RE power plant supplies a timely transmission of electricity, the project company receives payment for the power products from the power supplier. The renewable energy project BOT structure is shown in Fig. 1. In addition, the project company can apply for government authorized special renewable energy funds or can apply for financial subsidies. This additional income (including government subsidies) can assist the project company to cover the costs of the renewable energy plants or to repay bank loans or syndicated loan principal and interest.

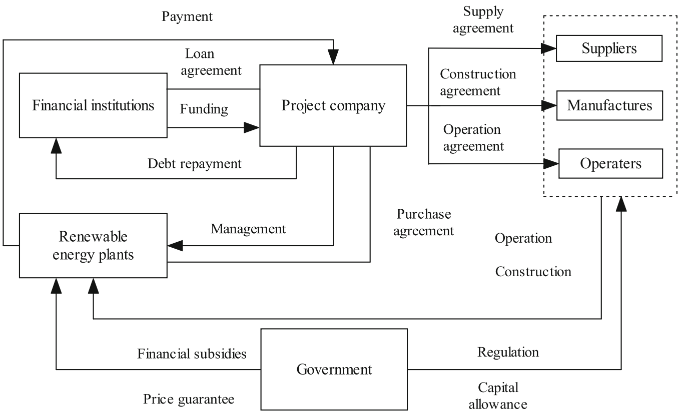


Fig. 1. BOT structure for renewable energy project

4.2 Financing Leases

Financing leases are a significant financing means for the development of China's renewable energy sector. Large scale renewable energy projects are less attractive because of the massive initial investment needed and the uncertainty risks. Lease financing programs refer to the relationship between the enterprises selling the RE equipment and a financial institution, for which the financial institution provides financing for the purchase of the equipment. The typical structure is depicted in Fig. 2. Finance leases for large-scale renewable energy power generation projects are initiated through commercial capital leases for the purchase of power generating equipment and technology, which is then recouped from elec-

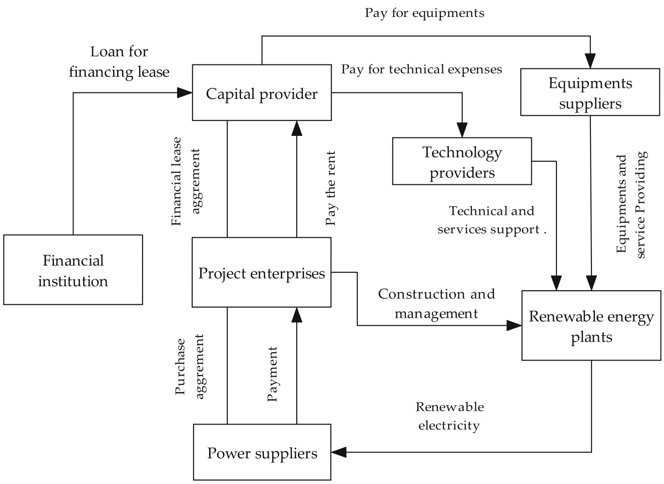


Fig. 2. The FL structure for an RE project

tricity users by the project operators to pay for the lease financing. In this way, a project sponsor can reduce their capital pressure as well as reduce the project's financial risk.

4.3 Third Party Financing

Third party financing (TPF) deals with customer sources, installation, engineering, maintenance and financing services for the RE system on the host customer's properties. The advantages of TPF contracts are that the source of funds is guaranteed and a number of services are provided which minimize commitment risk. TPF can effectively aid in the development of RE technologies in developing countries and rural areas with less investment. The third-party ownership model is shown in Fig. 3.

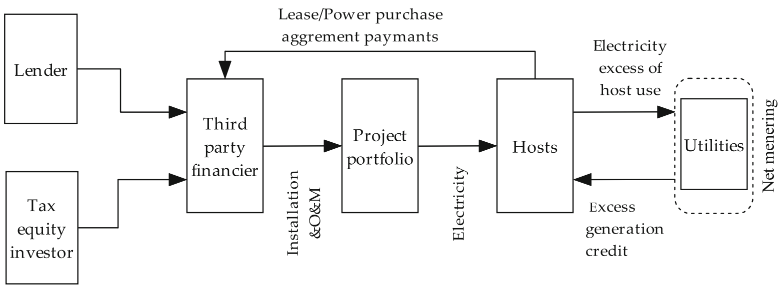


Fig. 3. Third-party ownership model

4.4 Venture Capital

Venture capital or private equity investment facilitates the provision of funding to firms in the industry that may have unproven or high risk technologies. From an investment point of view, almost all new energy projects have a corresponding capital risk, such as solar energy and wind energy related technological research and development projects. Venture capital generally involves a private placement of funds to assist in the establishment of emerging small and medium-sized RE enterprises which tend to have high *risk/high* return characteristics. At present, China's venture capital investment in new energy sectors is still in a trial stage; therefore, the investment is more cautious and the investment funds are often quite limited.

5 Issues of Renewable Energy Finance

This paper has closely examined optimization and innovation in RE finance issues to ensure RE sustainable development in China. From this close examination, there are several outstanding issues that need to be resolved to guarantee the continued long-term sustainable success of RE in China. The main issues are as follows:

(1) Incomplete financing systems for RE projects. Generally, investment options have focused on the industrialization of RE technologies only rather than on innovation. Second, domestic enterprises are not investing enough and have not been sufficiently integrated into RE development chains. However, for the RE true sustainable development, investment chains for renewable industry sectors and investment integration is necessary to ensure that the investments contribute to overall long-term sustainability.
(2) Lack of policy coordination and consistency. As there appear to be differences in the goals of each of the RE sectors, it has been difficult to ensure consistency in policy frameworks making it difficult to develop a long-term and effective policy system to support renewable energy"s sustainable development. Further, as energy issues encompass federal, state and local governments and several different departments such as energy, science and technology and agriculture, there has been inconsistency and a lack of coordination in policy development across these boundaries, ultimately eroding state macro-control and resulting in wasted investments due to duplication.
(3) Weakness Financial subsidies for renewable energy investment projects are much lower than for other sectors and China's support for renewable energy through taxation methods is fairly weak, with the actual taxation benefits for most renewable energy projects being similar to or of less benefit than conventional energy tax benefits.
(4) Inadequate investment in technical R&D and weak independent RE development innovation. At the moment, China lacks a powerful platform to support technological research. What is needed is a clear long-term continuous and rolling R&D development investment plan, with guaranteed funding to support R&D. Most core technologies in most domestic enterprises come from abroad, which will ultimately affect the sustainable development of the RE sector.

6 Conclusion

Investment, financing and corresponding policies have played a significant role in sustainable RE development. Investment, financing, finance means and related laws, regulations and major incentive mechanisms in China have progressed significantly in recent years. Various incentive mechanisms such as the RE preferential fiscal policies and R&D fund support for the commercialization of technologies has gone some way to decreasing the development costs of renewable energy projects. In the future, joint financing and policies and support mechanisms for renewable power generation need to be further strengthened. Fiscal and tax incentives have significantly relieved the financial burdens on renewable energy power generation enterprises and preferential tax policies for VAT, income tax and import duties for the renewable energy power generation industry as well as strong financial subsidies have provided invaluable support for renewable power projects. Tariff incentives, in particular, have ensured reasonable profits for renewable energy power generation enterprises. However, investment and financing renewable energy policies need to be strengthened further as China's RE financing methods and development policies have some inherent disadvantages. For the future, it is necessary to develop a range of coordinated strategies to ensure long-term efficient developments for the RE industry sector. In brief, this paper provides a comprehensive analysis of China's RE sustainable development. The major financing methods and identified incentive strategies and approaches can assist researchers and industry practitioners gain a better understanding of how RE sustainable development can be promoted by the Chinese government. The lessons learned from this China study can also be compared with or used to formulate innovative financing methods and incentive polices for RE sustainable development in other countries.

Acknowledgements. This research is supported by the Program of National Natural Science Foundation of China under Grant (No. 71503216) and (No. 71401136), the Social Science Planning Project of Sichuan Province under Grant (No. SC15C051), the Central Universities Project under Grant (No. 2016NZYQN13), the Technology Finance and Entrepreneurship Financial Research Center Project under Grant (No. JR201507), and the System Science and Enterprise Development Research Center Project under Grant (No. xq15c02).

References

1. Abolhosseini S, Heshmati A (2014) The main support mechanisms to finance renewable energy development. Renew Sustain Energy Rev 40:876–885
2. Bobinaite V, Tarvydas D (2014) Financing instruments and channels for the increasing production and consumption of renewable energy: Lithuanian case. Renew Sustain Energy Rev 38:536–545
3. Cao X, Kleit A, Liu C (2016) Why invest in wind energy? Career incentives and chinese renewable energy politics. Energy Policy 99:120–131
4. Center NRE (2016) The national renewable energy development report
5. Geng W, Ming Z et al (2016) China's new energy development: status, constraints and reforms. Renew Sustain Energy Rev 5:885–896
6. Koseoglu NM, van den Bergh JC, Lacerda JS (2013) Allocating subsidies to R&D or to market applications of renewable energy? Balance and geographical relevance. Energy Sustain Dev 17:536–545
7. Ming Z, Song X et al (2013) New energy bases and sustainable development in China: a review. Renew Sustain Energy Rev 20:169–185
8. Ng TH, Tao JY (2016) Bond financing for renewable energy in Asia. Energy Policy 95:509–517
9. Shen J, Luo C (2015) Renewable and sustainable energy reviews. Overall review of renewable energy subsidy policies in China—Contradictions of intentions and effects. Renew Sustain Energy Rev 41:1478–1488
10. Yang XJ, Hu H et al (2016) China's renewable energy goals by 2050. Environ Dev 20:83–90
11. Zhao ZY, Zuo J et al (2011) Impacts of renewable energy regulations on the structure of power generation in China—a critical analysis. Renew Energy 36:24–30

The Study on Factors Affecting the Attraction of Microblog – An Empirical Study Based on Sina Microblogging Site

Yue He, Yue Zhang(✉), Min Xiao, and Lingxi Song

Business School, Sichuan University, Chengdu 610064,
People's Republic of China
luner.zhang@foxmail.com

Abstract. New media like microblogging site becomes increasingly popular in China nowadays. Microblogging site has become an important social platform for information communication, and also an important marketing channel for enterprises. However, marketing messages are often submerged in a large amount of information on microblogging sites, so how to get more attention for their marketing microblogs becomes a concern of enterprises. Based on the data from Sina Weibo, we put forward "Attraction Index of Microblog" to measure how much attention a microblog has attract. Then we used several quantitative methods like correlation analysis and association rules to find the factors affecting the attraction of microblog, in order to provide some guiding ideas for marketing users to improve the attraction of their microblogs. With empirical studies, we found that microblogs published by authenticated users are more attractive. Whether the microblog has media information does not affect the attraction of microblog. The more fans and microblogs an authenticated user has; the more attractive the user's microblogs are. For a non-authenticated user, the number of fans somehow affects the attraction of their microblogs. The microblogs which are published at 7–10 am or 19–21 pm are more attractive.

Keywords: Microblog marketing · The attraction index of microblog · Influence factor · Association rules

1 Introduction

Sharing and transferring information, from useful knowledge to gossip, is an essential part of social life. In modern society, this needs are met by the social media, which allow people to share information more easily and rapidly. Microblogging site has gained tremendous popularity since their introduction in 2005, with the number of registered users growing exponentially [12]. Twitter, a microblogging service, is one of the most popular social network services [11].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_45

However, in China, Sina Weibo, similar with Twitter, is currently the most widely used microblogging platform with the most number of users. As of June 2016, the size of micro-blog users was 242 million, and the utilization rate is 34%, which has been gradually recovered compared with the end of 2015 [4]. The widespread use of Sina Weibo has attracted enterprises to pay attention to microblog marketing. Enterprises need to use microblog for branding promotion and products marketing.

With the rapid development of social network based on Web2.0, micro-blog services as an important social platform, which has the advantage of information sharing has attracted the participation of many network users. With the rapid development of Internet and wide application in many fields, most of the enterprises have begun to use the Internet for marketing activities of enterprises, micro-blog can focus on his fans and fans, which can realize real-time information sharing, business information dissemination. As an independent media on social media, the commercial value of micro-blog's business from the micro-blog platform, a large number of users and huge traffic. Enterprises can use micro-blog platform, and according to their own characteristics to implement a variety of marketing activities. Therefore, the paper hopes to find what kind of micro-blog is more likely to get people's attention, and help enterprises to make better social marketing.

2 Related Work

Since Twitter become people's important social networking platform and the important marketing channel of enterprises, there is a growing body of research on Twitter. The characteristics of Twitter and its applications are studied in various fields. Kwak et al. [18] studied the topological features of Twitter, and indicated that Twitter serves more as a news medium than a social network. Fischer and Reuber [5] use an inductive, theory-building methodology to develop propositions regarding how effectuation processes are impacted when entrepreneurs adopt Twitter. Chen [3] studied how active Twitter use gratifies a need to connect with others and the characteristics of user behavior.

There are many scholars have studied the user behavior characteristics on social network like Twitter, and also studied the effect of user behavior on user influence, information spread and information credibility. Cha et al. [1,2] studied the effects of user behavior characteristics such as number of fans and users' authoritativeness on microblog diffusion, and found that microblogs of opinion leaders, celebrities, politicians and entrepreneurs are more widespread. Räbiger and Spiliopoulou [16] proposed a complete framework for supervised separation between influential and non-influential users in a social network, finding that there are predictive properties associated with the activity level of users and their involvement in communities. These studies are focus on the influence of user identity (such as user occupation and user's authoritativeness) and user behavior characteristics (such as number of fans and contact degree between users and fans). They discuss about few influence of contents of tweet. Besides, user influence, information credibility and information spread effect have a certain correlation with popularity or attraction of message, but they are not all the same. Messages of influential users or of credibility may be not attractive. And attractive messages may not have good spread effect (i.e., a tweet receives many comments and likes, but fewer users' retweeting. We may say this tweet is attractive, but with less spread effect).

The largest microblogging website in China, Sina Weibo, born in 2009, also attracts much attention from researchers. In a comparison of Twitter and Sina Weibo, Gao et al. [7] analyzed the textual features, topics and sentiment polarities of posts for two microblogging websites, revealing significant differences between them. Liu et al. [14] analyzed the spatial distribution pattern of Chinese tourism by microblog, and revealed the factors which affect the spatial distribution of tourism. Gao Jie [6] makes some discussion on the celebrity microblog advertisement in Sina micro-blog. Combined with the features of government microblog, Xie et al. [13] constructs a theoretical model, and use structured questionnaire to validate the model. Yu et al. [17] examined the key trending topics on Sina Weibo, compared them with the observations on Twitter. They found that, on Sina Weibo, trends emerge almost completely attributed to reposts of entertainment content such as jokes and images, while on Twitter, the trends are always due to current global events and news stories.

There is also some research about user behavior of popular microblogs on microblogging sites. Qiu [9] qualitatively analyzed the influencing factors on popular microblog from the five aspects of disseminator, dissemination contents, audiences, dissemination channels and dissemination skills. Liu et al. [15] found that source trustworthiness, source expertise, source attractiveness, and the number of multimedia have significant effects on the information retweeting. Guan, Gao and Yang [8] found that male users are more likely to be involved in hot events, and messages that contain pictures and those posted by verified users are more likely to be retweet, while those with URLs are less likely. Lun et al. [10] found that retweeting and commenting are distinct types of microblogging behaviors. Retweeting aims to disseminate information in which the source credibility and posts' in formativeness play important roles, whereas commenting emphasizes social interaction and conversation in which users' experience and posts' topics are more important. Most of them think that a popular microblog is retweeted many times or commented by many people, few of them consider the number of "likes" as the factor to measure the popularities of the microblog.

On the basis of previous studies, this paper combine the forwarding number, the number of comments and points like the number of calculation of "micro-blog attention" score to measure micro-blog attention degree. And through correlation analysis, association analysis to analyze the factors affecting micro-blog attention.

3 Data Collection and Methods

3.1 Data Collection

We randomly extracted 107,328 original microblogs from Sina Weibo with the LocoySpider 8.0. The microblogs were posted from March to April 2015. The data include information about the user (including user's nickname, the number of fans, followers (who the user follows) and microblogs, whether the user is authenticated) and microblog (including microblog content, the number of people who like it, the number of comments, retweeting times, whether with multimedia information, posting time).

Because the data has different dimension or different order of magnitude, which will affect the comparison and analysis of the data, so data standardization is needed first. Data is standardized between the interval [0, 1] with the SoftMax standardization method. The formula to standardize x_i to $x_{i\text{new}}$ is as follows:

$$x_{i\text{new}} = \frac{1}{1 + e^{-x_t}}. \tag{1}$$

The x_t in the formulae (1) is calculated as:

$$x_t = \frac{x_i - \bar{x}}{\sigma}. \tag{2}$$

3.2 MAI ("Attraction Index of Microblog")

Liu et al. [15] measured the attraction of microblog with retweeting times. Lun et al. [8] used number of comments to measure microblog attraction. On Sina Weibo, users can not only "retweet" and "comment" a microblog, but also "like" a microblog. If a user agrees with or appreciates a microblog, he/she may push the button "like". A microblog receives more "like", which means the microblog attracts more attention and love. Therefore, we believe that "like" can also reflect the microblog attention. So the number of "like" should be added into "Attraction Index of Microblog" model.

Compared with "like" and "comment", retweeting is more unique in the three kinds of user behaviors. Retweeting show that a user not only likes this microblog, but is more willing to share with more people. Microblog through retweeting layer upon layer, its mode of transmission is fission, which leads to exponential growth of the number of users it attracts. Therefore, the retweeting times in "Attraction Index of Microblog" model is in the form of exponential function.

According to the characteristics of these three parameters, the number of "like" and comments and the retweeting times should be combined to evaluation microblog attraction. So "Attraction Index of Microblog" was built:

$$\text{MAI} = (\alpha ML_i + \beta MC_i)\, e^{MF_i}. \tag{3}$$

ML_i, MC_i and MF_i respectively represent the number of "likes", comments and retweeting times which are standardized. α, β respectively represent the coefficients of the number of “likes” and comments. In order to determine the coefficient α, β, we used the method of partial correlation analysis. Based on the microblog data, we calculated the partial correlation coefficients of the number of “likes” and comments with retweeting times respectively, which are as the value of α, β.

Partial correlation analysis [10] is used to analyze the partial correlation coefficients of two variables among multiple variables under the control of the effect of other variables, in order to represent the degree of linear relationship.

Suppose there are x_1, x_2, x_3 three variables, it is required to calculate the partial correlation coefficients between the variables x_1 and x_2, with the excluding the impact of variable x_3. The partial correlation coefficient is denoted as r12,3:

$$r_{12,3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{13}^2\sqrt{1 - r_{23}^2}}}. \tag{4}$$

In formula (4), r_{ij} represents the simple correlation coefficient of variable x_i and x_j, and $i, j = 1, 2, 3$.

The results of partial correlation analysis with SPSS19.0 are shown in Table 2

Table 1. The partial correlation coefficients of the number of “likes” and comments with retweeting times respectively.

	Number of “likes”	Number of comments
Retweeting times	0.457	0.523

Table.1 shows that the coefficients of the number of “likes” $\alpha = 0.457$ and the coefficients of the number of comments $\beta = 0.52$.

Eventually, “Attraction Index of Microblog” formula was:

$$\mathrm{MAI} = (0.457ML_i + 0.523MC_i)\, e^{MF_i}. \tag{5}$$

3.3 Association Rules

Microblog with pictures, video and other multimedia information will bring more visual or auditory stimuli, so these kinds of microblogs more attractive. If the user is authenticated, whose microblog are more reliable and authoritative. So, it is more likely to be commented or retweeted by others, the microblog attraction may also be higher.

We used Apriori model in Clementine 12.0 to analyze the relationship between MAI and microblog multimedia information, users’ authoritativeness respectively. We set the support level with 5%, and the confidence level with 10%.

Association rules is used to find out the relevance hidden in the database, and expressed in the form of rules. It is shaped like an implication "$X \Rightarrow Y$". The associational rule refers to the sets of attribute-values, which frequently appeared in data set recognition, also named as frequent item-sets. Using these frequent sets, association relation rules process is set up. In association rules, count (X) is the number of tuples that object set D contains set X. Support for set X is:

$$\text{Support}\,(X) = \frac{\text{count}\,(X)}{|D|}. \tag{6}$$

Confidence is the proportion of the number of tuples that contains set X and Y and the number of tuples that contains set X:

$$\text{Confidence}\,(X \Rightarrow Y) = \frac{\text{count}\,(X \cap Y)}{\text{count}\,(X)}. \tag{7}$$

Apriori algorithm proposed by Agrawal in 1994 is a relatively simple and general algorithm. The algorithm has two key steps: the first is to find all frequent item sets, and the second is to generate strong association rules.

4 Results

4.1 Descriptive Statistics of MA

For the 107328 microblogs, according to the formula (5) we calculated each microblog's MAI. To limit the value of MAI within [0, 1], we used formula (8) to standardize the value of MAI, then the final MAI of each microblog was got.

$$x_{\text{new}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}. \tag{8}$$

The average of MAI is 6.62×10^{-3}. The minimum value is 0, and the maximum value is 1. The MAI of most microblogs was low. The number of microblogs whose MAI is equal to 0 is 65270, accounting for 60.8% of the total number of microblogs. The number of microblogs whose MAI is greater than the average MAI is 14726, accounting for 13.7% of the total number of microblogs. The MAI distribution is shown in Fig. 1. Due to the small value of MAI, we expanded the value 100 times to display (because the microblogs whose MAI greater than 0.1 is few, Fig. 1 is not in full).

According to the MAI, microblogs were divided into two categories: attractive microblogs and unattractive microblogs. Microblogs whose MAI is greater than the average MAI are attractive microblogs, while Microblogs whose MAI is less than the average MAI are unattractive microblogs. The following analysis is based on this.

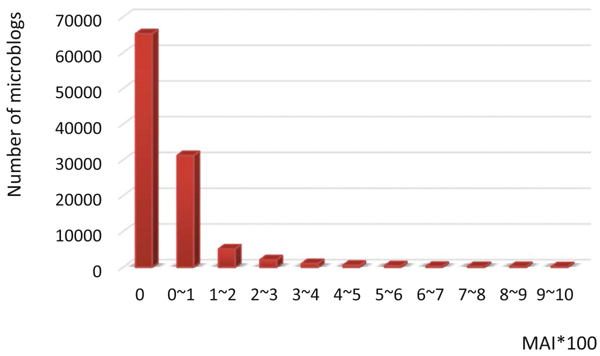


Fig. 1. The MAI distribution situation

4.2 Microblog Multimedia Information and User Authoritativeness

The results of association rules analysis is shown in Table 2:

Table 2 shows:

- Whether the microblog with multimedia information does not affect the microblog attraction.
- Whether the user is authenticated has effect on the microblog attraction. The microblogs posted by authenticated users are more attractive than those posted by non-authenticated users.

Table 2. The extraction results of association rules of MAI

Y	X	support (X)	confidence $(X \Rightarrow Y)$
With multimedia	Attractive	13.72%	72.36%
Attractive	With multimedia	74.52%	13.43%
Authenticated user	Attractive	13.72%	21.49%
Attractive	Authenticated user	8.69%	34.21%

The distribution of MAI of the microblogs respectively post by authenticated users and non-authenticated users is shown in Fig. 2. As we can see, the MAI of microblogs posted by non-authenticated users are lower the authenticated users. It also confirmed whether users' authoritativeness is a key factor affecting the microblog attraction.

4.3 Information of User Behavior

The Spearman correlation coefficients between MAI and the number of user's fans, followees and microblogs respectively are shown in Table 3.

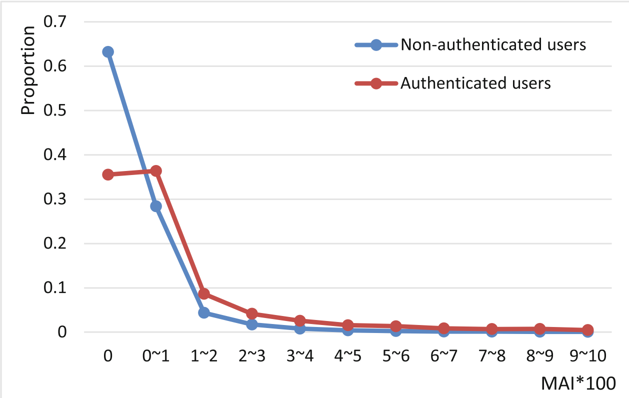


Fig. 2. The distribution of MAI of the microblogs respectively post by authenticated users and non-authenticated users

Table 3. The Spearman correlation coefficients between MAI and the number of user's fans, followers and microblogs respectively

The Spearman correlation coefficient		N_fan[a]	N_fol[b]	N_mcb[c]
MAI	Authenticated user	0.522**	0.113	0.241**
	Non-authenticated user	0.271**	0.074	0.097

[a]: N_fan: The number of fans. [b]: N_fol: The number of followees. [c]: N_mcb: The number of microblogs. **: Correlation is significant at the 0.01 level (2-tailed).

Table 3 shows that for different users, the effect of user behavior on microblog attraction is different. For authenticated users, the number of fans have a great effect on the microblog attraction, with the correlation coefficient reached 0.522, followed by the number of microblogs, and the number of followees has little effect. For non-authenticated users, the number of fans has less effect on the microblog attraction, and the number of followees and microblogs has almost no effect.

4.4 Posting Time

Generally speaking, most people's leisure time is the two period of time at noon and in the evening, so we can guess that microblogs posted in the two period of time are more than other time. So, how about attractive microblogs? Are there also some regularities the posting time of attractive microblogs follow?

In order to study the effects of microblog posting time on microblog attraction, we analyzed the posting time of attractive microblogs. We calculated the number of total microblogs and attractive microblogs per hour, and compared the two distributions of posting time of total microblogs and attractive microblogs.

Due to the large gap of the absolute number of total microblogs and attractive microblogs, it is not suitable to compare the absolute value directly. So, we calculated the proportion of total microblogs and attractive microblogs per hour respectively. For example, the proportion of total microblogs per hour is calculated by dividing the number of total microblogs per hour by the number of all microblogs (107328), so as the proportion of attractive microblogs. The posting time distribution of total microblogs and attractive microblogs is shown in Fig. 3.

As Fig. 3 shows, the posting time distribution of total microblogs conforms to the guess that microblogs posted at noon (11:00–13:00) and in the evening (21:00–23:00) are more than that posted at other time. And there are few microblogs posted between 1:00–7:00 am. We can also find that attractive microblogs posted between 8:00–11:00 and 19:00–21:00 are more than that at other time, and the fluctuation is more obvious.

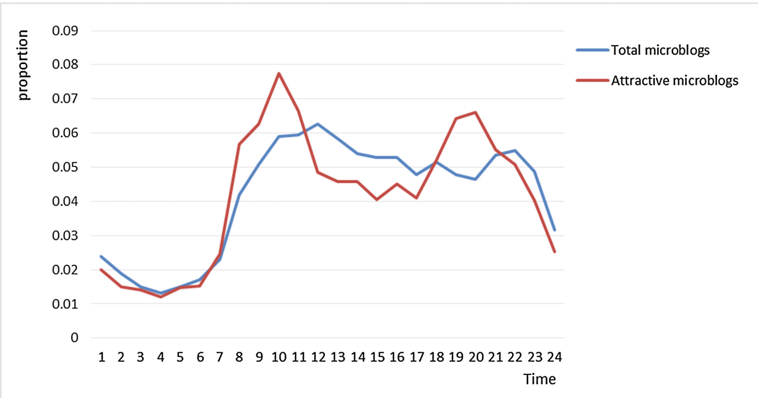


Fig. 3. Posting time distribution of total microblogs and attractive microblogs

In order to observe the posting time distribution of attractive microblogs vividly, the number of attractive microblogs per hour was divided by the number of total microblogs per hour. The proportion of them is shown in Fig. 4, which will eliminate the interference of regularities of the posting time distribution of total microblogs.

From Fig. 4, we can obviously see that the microblogs posted at 8:00–11:00 and 19:00–21:00 are more attractive. Besides, comparing the posting time of total microblogs and attractive microblogs, we can find that attractive microblogs are posted earlier about two hours than total microblogs. According to our real-life, we can explain that most people at the leisure time (11:00–13:00 and 21:00–23:00) are likely to see the microblogs posted 2 or 3 h in advance. So, these microblogs are more likely to be read, increasing the possibility to be retweeted/comment/liked, resulting in becoming attractive microblogs.

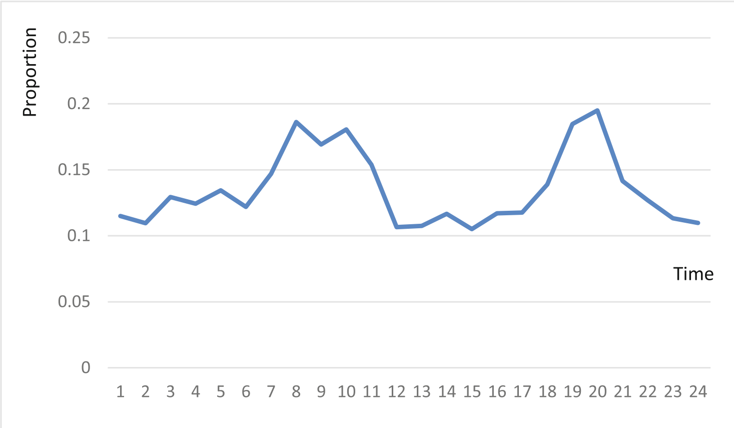


Fig. 4. Proportion of the attractive microblogs within all microblogs per hour

5 Summary and Discussion

Through the empirical research, we found the factors affecting the microblog attraction as follows:

- Users' authoritativeness. If the user is authenticated, he/she is more credible and popular, so his/her microblogs also can obtain others' trust and concern, which makes the microblogs attractive.
- The information of user behavior. For the authenticated users, factors affecting the microblog attraction is the number of fans, followed by the number of microblogs. The more the number of fans or microblogs, the stronger the microblog attraction is. For non-authenticated users, factors affecting the microblog attraction is the number of fans, but it does not affect the microblog attraction as much as the number of authenticated users' fans does.
- Microblog posting time. The study shows that attractive microblogs are mostly posted between 8–11 a.m. and 7–9 p.m., so microblogs posted in these two periods of time would attract more attention.

Our study possesses certain innovation and pertinence. We put forward "Attraction Index of Microblog" to definitely measure how much attention the microblog has attract. And we used several quantitative data analysis methods to analyze how these factors affect the microblog attraction. The findings of our study also have some practical implications. Our study provides some guiding ideas for Weibo users to improve the attraction of their microblogs. Weibo users especially advertisers who aim to market their product and disseminate their corporate culture through microblogging sites, for example, must try to obtain authentication to attract more fans. And they must keep posting microblogs to provide enough messages. It is not necessary to insert media information like pictures or videos into microblogs. It is also important to post microblog in the two period of time, which can make the microblog be seen by more people.

6 Limitations and Directions for Future Studies

This paper has certain innovation, but there are still shortcomings and limitations, which are worthy of exploration and improvement for future research, mainly in the following points:

- We only studied the effects of microblog post form and the user information to the microblog attraction, not considering the influence of microblog content. So in the following research, we can continue to study the effects of microblog content to its attraction, including the microblog sentiment, and topic types etc.
- For the analysis of each factor affecting the microblog attraction, we only used a method. However, different methods as the principle or perspective is different, may lead to different results. Therefore, in the future study, we can analyze the same problem with different methods, and compare the difference of results to choose the more appropriate methods and results.
- The current study examined microblogging behavior using Sina Weibo in the Chinese context. Other social networking sites like Twitter, Facebook have much similarity to microblogging site, but in many ways, they are different, such as different technology, different culture, and different user behaviors between different platforms. Therefore, it would be unacceptably risky to conclude that the findings of this study represent a cross-contextual evaluation of microblogging behavior. Future researchers should make extensive efforts to explore the difference of user behavior between different social networking platforms.

Acknowlegement. The authors acknowledge the financial support National Natural Science Foundation Project of China, project (No. 71372189), and Sichuan University Fundamental Research Funds for the Central Universities Project (No. skqy201406). The author is grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

1. Cha M, Haddadi H et al (2010) Measuring user influence in twitter: the million follower fallacy. In: International conference on weblogs and social media, ICWSM 2010, Washington, DC, USA, May 2010
2. Cha M, Benevenuto F et al (2012) The world of connections and information flow in Twitter. IEEE Trans Syst Man Cybern Part A Syst Hum 42(4):991–998
3. Chen GM (2011) Tweet this: a uses and gratifications perspective on how active twitter use gratifies a need to connect with others. Comput Hum Behav 27(2):755–762
4. CNNIC (2016) The 38th statistical report on internet development in China. Technical report. http://www.cnnic.net.cn/ (in Chinese)
5. Fischer E, Reuber AR (2011) Social interaction via new social media: (how) can interactions on twitter affect effectual thinking and behavior? J Bus Ventur 26(1):1–18

6. Gao J (2016) Research on the advertising value of celebrity micro-blog. PhD thesis, Jilin University, Changchun (in Chinese)
7. Gao Q, Abel F et al (2012) A comparative study of users' microblogging behavior on Sina Weibo and Twitter. Springer, Heidelberg
8. Guan W, Gao H et al (2014) Analyzing user behavior of the micro-blogging website Sina Weibo during hot social events. Phys A Stat Mech Appl 395(4):340–351
9. Jin Q (2013) An analysis of the influencing factors on popular micro-blog. PhD thesis, Zhengzhou University (in Chinese)
10. Kim JM, Jung YS et al (2011) Partial correlation with copula modeling. Comput Stat Data Anal 55(3):1357–1366
11. Ko J, Kwon HW et al (2014) Model for Twitter dynamics: public attention and time series of tweeting. Phys A Stat Mech Appl 404(24):142C149
12. Li F, Du TC (2014) Listen to me - evaluating the influence of micro-blogs. Decis Support Syst 62(2):119–130
13. Liang CC (2015) Factors influencing office-workers' purchase intention though social media: an empirical study. Int J Customer Relat Mark Manage 5:109–116 (in Chinese)
14. Liu DJ, Jing HU et al (2015) Spatial pattern and influencing factors of tourism micro-blogs in China: a case of tourism Sina micro-blogs. Sci Geogr Sin 35(6):717–724
15. Liu Z, Liu L, Li H (2012) Determinants of information retweeting in microblogging. Internet Res 22(4):443–466
16. Räbiger S, Spiliopoulou M (2015) A framework for validating the merit of properties that predict the influence of a Twitter user. Expert Syst Appl 42(5):2824–2834
17. Yu L, Asur S, Huberman BA (2011) What trends in Chinese social media? arXiv preprint arXiv:107:3522
18. Yu S (2014) Microblogging becomes the first listed Chinese social media. Accessed People's daily: http://tech.sina.com.cn/i/2014-04-18/09309329858.shtml

A Genetic Algorithm Approach for Loading Cells with Flow Shop Configuration

Patrick Gannon and Gürsel A. Süer(✉)

Department of Industrial and Systems Engineering,
Ohio University, Athens, OH 45701, USA
suer@ohio.edu

Abstract. This paper proposes a three-phase methodology for worker allocation and flowshop scheduling in a multistage manufacturing environment. A case study using a shoe manufacturing plant is examined. The proposed methodology consists of genetic algorithm and mathematical models. Phase 1 allocates workers in between operations in two separate stages in the manufacturing process. Phase 1 evaluates how workers are allocated between and within the manufacturing stages. Phase 2 uses worker allocations from phase 1 to perform cell loading based on machine-level-based similarity using genetic algorithms. Four different Genetic Algorithm approaches are proposed and evaluated. Furthermore, the impact of different manpower levels are also studied. Phase 3 schedules products within cells to optimize makespan, total tardiness, or number of tardy jobs.

Keywords: Genetic algorithms · Cell loading · Flowshop scheduling · Worker allocation

1 Introduction

The goal of this paper is to provide a methodology to most effectively utilize available manpower and optimize the scheduling of products with regards to multiple performance measures. The proper utilization of resources is crucial to success and survivability of any company and the proposed methodology seeks to improve utilization of resources in multi-stage labor-intensive manufacturing environments. Using a shoe manufacturing plant as a case study, this paper addresses the problems of manpower allocation, cell loading, and flowshop scheduling. The shoe manufacturing plant has several cell groups, consisting of a lasting cell (LC), a rotary machine cell (RMC), and a finishing cell (FC) as shown in Fig. 1. The LC and FC are labor intensive manufacturing cells consisting of several sequential operations. A simplified version of the RMC is used in this paper, so the majority of the focus of this paper is on the LC and FC. There are fewer operation types in the LC and FC than there are workers available,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_46

so several operators can be assigned to each operation. With correct manpower allocation, a balance can be achieved and the cells' output can be maximized. Considering the worker allocation and capacity requirements cell loading can be performed. By carefully assigning products to cells, the benefits of cellular manufacturing, such as reductions in flowtime, setup times, and work-in-progress inventories, can be realized. Once products are assigned to cells, scheduling is performed in each cell group independently. Scheduling is necessary when resources are limited and need to be optimally utilized for several competing tasks.

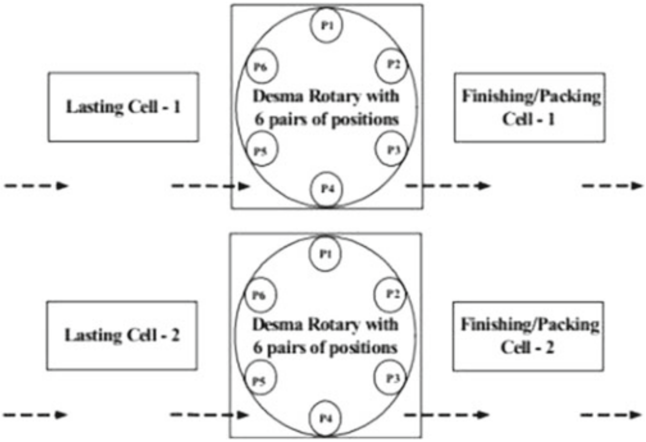


Fig. 1. Shoe manufacturing system

1.1 Products

This plant produces shoes with a variety of designs, colors, and materials. The shoes are available in two sole types, full shot or midsole. Of these two sole types, the shoes can be made using either PVC or TPR and are available in three colors (black, honey, and nicotine). The breakdown of this product structure is shown in Fig. 2.

The demand for shoes is based on their size. The sizes range from size 5 to size 15 and demand is greatest closest to middle of the size range. The demand curve shown in Fig. 3 demonstrates the relationship between size and demand.

1.2 Rotary Machine

The Rotary Machine Cell (RMC) is the bottleneck of the manufacturing cell. The RMC has six available positions, each of which can process a pair of shoes. The Rotary Machine also has a tank to hold the material to be injected. Unprocessed shoes are loaded into the machine and injected with either PVC or TPR in order

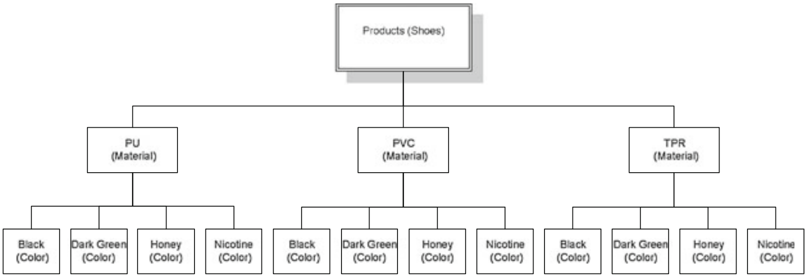


Fig. 2. Product types

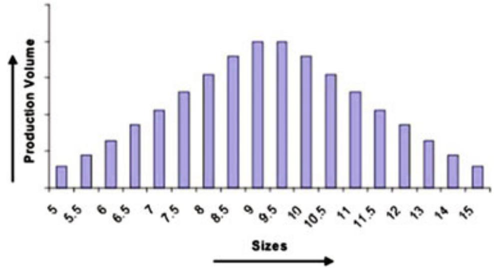


Fig. 3. Typical shoe demand by size

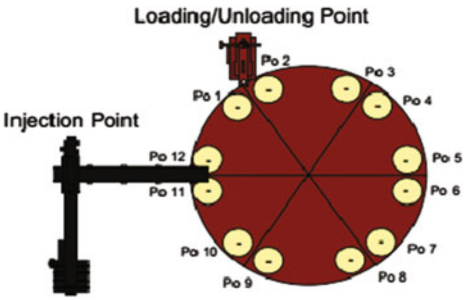


Fig. 4. Rotary injection molding machine

to form the soles of the shoes. A setup period is needed to switch materials or colors to be injected since the machine possesses only one tank to hold the material.

The Rotary Machine consists of a loading/unloading station, an injection station, and four cooling stations. Figure 4 illustrates the setup of the rotary machine. A shoe's processing time on the Rotary Machine is influenced by the mold type (full shot or mid-sole), the size of the shoe, and whether it is designed for males or females. The average injection molding time for full shot shoes is 0.356 min for male shoes and 0.3 min for female shoes. Midsole shoes require less processing, having average injection molding times of .286 min for male shoes and .25 min for female shoes. Table 1 shows injection times for the various designs at the available sizes. Times are classified as Fullshot (FS), Midsole (MS), male (M) and Female (F).

Table 1. Injection Times

FS/MS	M/F	Different sizes								
		1	2	3	4	5	6	7	8	9
		5	5.5	6	6.5	7	7.5	8	8.5	9
FS	M	-	0.224	0.247	0.269	0.281	0.289	0.290	0.301	0.306
FS	F	0.154	0.207	0.216	0.229	0.23	0.237	0.308	0.338	0.356
MS	M	-	-	0.160	0.162	0.203	0.239	0.24	0.242	0.245
MS	F	0.109	0.116	0.183	0.190	0.196	0.260	0.261	0.290	0.305
FS/MS	M/F	Different sizes								
		10	11	12	13	14	15	16	17	18
		9.5	10	10.5	11	11.5	12	13	14	15
FS	M	0.32	0.342	0.372	0.426	0.428	0.448	0.460	0.499	0.556
FS	F	0.368	0.395	-	0.424	-	0.440	-	-	-
MS	M	0.247	0.273	0.293	0.311	0.349	0.398	0.401	0.405	0.415
MS	F	0.31	0.342	-	0.344	-	0.347	-	-	-

Shoes loaded into the rotary machine cannot be unloaded until all shoes in the machine have been processed. This leads to shoes with lower processing times spending more extra time in the machine relative to shoes with higher processing times. This extra time leads to extra material being injected into their molds. Because of this, it is beneficial to minimize variations in the processing times of shoes that are to be loaded together. However, this problem has been addressed in prior works [6,12,16], so for the purpose of this paper, one size will be assigned at a time in the rotary machine.

1.3 Molds

Different molds are needed for processing shoes depending on size and type (Full Shot or Mid Sole). The molds carry a prohibitive cost, limiting their availability. As shoes of varying size and type are processed in the RMC, the molds have to be constantly changed in order to process the corresponding shoe. Also, the molds have to be cleaned when the material or color is changed. The problem considering a limited number of molds for each size has been addressed previously in the literature. For this study, it is assumed that there are no mold availability restrictions.

2 Literature Review

This paragraph will briefly discuss some prior cell loading and manpower decision work done in the literature. Süer, Saiz, Dagli, and Gonzalez [14] expanded research in the cellular control area to multi-cell environment, which necessitates cell loading considerations. Süer [13] developed a two-phase methodology to find optimal manpower allocation and cell loads simultaneously. An extension of this work was done by Süer [13]. This work add lot-splitting considerations to the two phase approach to find optimal manpower allocation and cell loads and found that lot splitting is advantageous when setup times are negligible. Akturk [1] addressed the production planning problem in cellular manufacturing systems. By using capacity constraints to evaluate the impact of cell loading decisions on lower levels, they were able to more accurately portray the cellular manufacturing system and its operation. Saad [9] developed a multi-objective optimization technique to load and schedule cellular manufacturing systems using simulation and a taboo search algorithm. Babayigit [2] analyzed the problems of manpower allocation and cell loading with a mathematical model and genetic algorithm, further extending previous works by Süer [13,14].

Süer [15] proposed a three-phase methodology to address cell loading and product sequencing in labor-intensive cells. Optimal manpower allocation and cell loads are found in the first two stages and the third stage treats product sequencing as a travel salesman problem to minimize intra-cell manpower transfers. Some scheduling work done in the literature is discussed in this paragraph. Nakamura [8] considered group scheduling on a single stage to minimize total tardiness and found that scheduling problems of moderate size can be ordered to reduce the number of schedules to be searched. Sinha [10] found that a desired level of throughput and optimum work-in-progress in a cell can be achieved through sequencing, reduced batch size, and period batch control. Gupta [5] developed an optimization algorithm and heuristic procedures to schedule jobs on two identical machines. The objective was to find a schedule with optimal total flowtime that gives smallest makespan. Kim [3] used a restricted taboo search algorithm to schedule jobs on parallel machines in order to minimize the maximum lateness of jobs. Gupta [4] studied single machine group scheduling with sequence-independent family setup times in order to minimize total tardiness. The heuristics developed were shown to be effective in minimizing total tardiness.

Subramanian [11] focused on cell loading and job scheduling at a shoe manufacturing plant. The objective of this research was to develop heuristics for cell loading and combine these with simple scheduling methodologies to load all jobs without exceeding capacity. This capacity was set using a MaxCap methodology, which is the maximum capacity to be loaded. This methodology is critical in determining the number of cells required and how the load is spread across different cells in the facility. An interesting finding from this research is that the optimal result from the cell loading stage may not remain optimal after the cell scheduling stage. Urs [7] developed three scheduling heuristics using basic machine scheduling methodology. The heuristics aimed to minimize makespan in the rotary machine scheduling. The best performing heuristic was the minimum difference in cycle time heuristic. This research stands out from prior works by assuming a limited number of molds for each size, meaning multiple sizes can be run on the Rotary Molding Machine simultaneously. Mese [7] studied cell loading and family scheduling in a cellular manufacturing environment. This study is distinguished from prior research by giving every job in a family an individual due date. A focus of this study was the tradeoff between meeting due dates of jobs and reducing total setup time. Setup times are reduced if all jobs of a family are scheduled together, but this could cause delays and tardiness in other jobs in other families. A way to offset this is to allow family splitting, but this increases setup times and in return may increase the number of tardy jobs. The author used mathematical modeling and a genetic algorithm to solve this problem. The mathematical model was slow and impractical to use for larger problems. The GA developed was both effective in finding optimal or near optimal solutions and efficient.

Süer [16] developed a three phase methodology to perform cell loading and scheduling in a shoe manufacturing company. This research used three family definitions (sub-families, families, and superfamilies) in the cell loading process. These different family definitions are used at different stages of the planning and scheduling process. Superfamily definition was used to determine the number of cells of each type. Family definitions are used after product-to-cell assignments. These families allow the number of setups to be minimized. Finally, the subfamily definition is used in the implementation of heuristics before cell loading begins. This research also used the MaxCap methodology used by Subramanian [12]. A valuable conclusion from this research is that the best post loading results did not give the best post scheduling results. This demonstrated the need to consider that isolating a single level of a multilevel problem may not result in the best solution for the overall problem. Only the injection molding cell is covered in this research, leaving further research to be done regarding the lasting and finishing cells in the shoe manufacturing company.

3 General Methodology

The three-phase methodology proposed to solve the problem is described in the following sections.

Phase 1 - Manpower Allocation

Thirty-five workers are split between the LC and FC. The six worker splits evaluated are 15/20, 16/19, 17/18, 18/17, 19/16, 20/15, where 15/20 represents 15 workers available for allocation in the LC and 20 available in the FC. The workers allocated to the LC and FC are then allocated among the operations within those cells. A mathematical model is used to optimally allocate manpower in order to maximize production rates. The model is run separately for each product at each worker level in both the LC and FC.

Phase 2 - Cell Loading

Products are assigned to cells in order to maximize the machine-level based similarity among products in the cells and also minimize the number of cells opened. A genetic algorithm is used to assign products to cells.

Phase 3 - Flowshop Scheduling

After jobs have been assigned to cells, flowshop scheduling is performed in each cell to minimize one of three performance measures. The performance measures considered are makespan, total tardiness, and number of tardy jobs. The methodology is summarized in Table 2.

Table 2. Summary of General Methodology

Hase	Method	Objective	Cells	Purpose
1-Worker allocation	Mathematical model	Maximize production rate	LC, FC	Assign workers to stations in LC & FC
2-Cell loading	Genetic algorithm	Maximize machine level similarity between products within cells	LC, FC	Assign products to cells, minimize machines required
3-Flowshop scheduling	Mathematical model	Minimize makespan, number of tardy products, total tardiness, number of machines	LC1-RC1-FC1	Determine sequence of LC2-RC2-FC2 products along with start LCn-RCn-FCn and completion time independently

4 Manpower Allocation Phase

In phase 1, workers are allocated to stations in the LC and FC, considering various worker levels, in order to maximize production rates. The model is solved for each product at each worker level and in both the LC and FC.

The index, parameters and decision variables are defined below:

Index:

j : Operation

Parameters:

R : Production rate for the product;

t_j : Unit operation time for operation j;

u_j : Maximum number of operators for operation j;

W : Total number of workers available

Variable:

m_j : Number of workers assigned to operation j

Objective Function:

$$\max \ Z = R \tag{1}$$

$$\text{s. t. } m_j \left(\frac{1}{t_j}\right) - R \geq 0, j = 1, 2, \cdots, s \tag{2}$$

$$m_j \leq U_j, j = 1, 2, \cdots, s \tag{3}$$

$$\sum_{j=1}^{s} m_j \leq W \tag{4}$$

m_j integer and positive for all j

R positive.

Equation (1) shows the objective function of the mathematical model, which is to maximize the production rate. The relationship between number of workers at a station and the production rate is determined in Eq. (2), ensuring enough workers are assigned to each operation to meet the desired production rate. Eq. (3) establishes an upper limit on the number of workers allowed at an operation. Equation (4) ensures that the total number of workers assigned to the stations does not exceed the total number of workers in the system. An example will be used to explain each phase of the problem. An example featuring 10 products each with five operations in both the LC and FC will be used. Unit processing times are given in Table 3. The mathematical model described earlier in this section was run using these processing times. The optimal worker allocations for product 1 at all worker levels are shown in Table 4. Table 5 shows optimal worker allocations for all 10 products at the 15/20 worker level.

Table 3. Unit processing times (minutes/unit) for the example problem

Product	Lasting cell					Finishing cell				
	Operations					Operations				
	1	2	3	4	5	1	2	3	4	5
1	1.41	1.36	0.76	0.65	0.39	0.44	0.41	1.49	0.54	1.15
2	0.80	0.78	0.85	0.76	0.59	1.23	1.10	0.32	0.63	1.17
3	0.79	1.45	1.13	1.98	0.53	1.07	0.43	0.91	1.21	1.12
4	0.31	0.54	2.79	0.66	0.95	0.77	0.99	0.95	1.47	0.33
5	1.26	0.43	1.39	1.28	1.40	0.58	0.68	1.49	0.28	0.40
6	0.73	0.57	1.04	0.78	0.94	2.18	0.43	0.42	1.72	0.23
7	0.57	0.78	1.23	0.19	0.36	0.60	0.44	1.12	0.81	1.00
8	0.43	1.42	1.07	0.25	0.79	1.64	2.05	0.70	0.31	2.55
9	0.87	0.61	1.04	0.87	0.68	0.48	0.59	0.26	0.15	0.41
10	1.98	0.62	1.25	1.25	0.80	1.39	2.10	0.62	0.63	1.17

Table 4. Worker allocation combinations for product 1

Product	Worker level	Lasting cell						Finishing cell					
		Operation					Prod. rate (U/min)	Operation					Prod. rate (U/min)
		1	2	3	4	5		1	2	3	4	5	
1	15/20	4	4	3	2	2	2.84	2	2	7	3	6	4.55
1	16/19	5	4	3	2	2	2.94	2	2	7	3	5	4.35
1	17/18	5	5	3	2	2	3.08	2	2	6	3	5	4.03
1	18/17	5	5	3	3	2	3.55	2	2	6	2	5	3.7
1	19/16	6	5	3	3	2	3.68	2	2	6	2	4	3.48
1	20/15	6	6	3	3	2	3.95	2	2	5	2	4	3.36

Table 5. Worker allocation for all products at 15/20 worker combination

Product	Worker level	Lasting cell						Finishing cell					
		Operation					Prod. rate (U/min)	Operation					Prod. rate (U/min)
		1	2	3	4	5		1	2	3	4	5	
1	15/20	4	4	3	2	2	2.8369	2	2	7	3	6	4.5455
2	15/20	3	3	3	3	3	3.5294	5	5	2	3	5	4.065
3	15/20	2	3	3	5	2	2.069	4	2	4	5	5	3.7383
4	15/20	1	2	7	2	3	2.509	4	4	4	6	2	4.0404
5	15/20	3	1	4	3	4	2.3256	3	4	8	2	3	5.1724
6	15/20	3	2	4	3	3	3.1915	8	2	2	7	1	3.6697
7	15/20	3	4	5	1	2	4.065	3	2	6	4	5	4.5455
8	15/20	2	5	4	1	3	3.5211	4	6	2	1	7	2.439
9	15/20	3	2	4	3	3	3.2797	5	6	3	2	4	9.7561
10	15/20	5	2	3	3	2	2.069	5	7	2	2	4	3.1746

The mathematical model was solved using ILOG OPL 6.3. The optimal production rates are determined in both the LC and FC at the given worker level. For example, product 2 at the 15/20 worker level for the Lasting Cell requires three workers at each operation and results in a production rate of 3.52 units/minute.

5 Cell Loading Methodology

Cell loading is performed using a genetic algorithm approach to maximize the similarity among products while minimizing the number cells opened. Similarity between jobs is measured using the machine level-based similarity coefficient between products i and k developed by Süer and Ortega [14]. The coefficient is calculated using Eq. (5). A sample calculation using the worker allocations in the lasting cell for products 1 and 2 is shown in Eq. (5).

$$\mathrm{MLB} - SC_{ik} = \frac{\sum_{j=1}^{s} \min(m_{ij}, m_{kj})}{\sum_{j=1}^{s} \max(m_{ij}, m_{kj})}, \tag{5}$$

$$\mathrm{MLB} - SC_{12} = \frac{\min(4,3) + \min(4,3) + \min(3,3) + \min(2,3) + \min(2,3)}{\max(4,3) + \max(4,3) + \max(3,3) + \max(2,3) + \max(2,3)}$$

$$= \frac{13}{17} = 0.76.$$

This calculation is performed for each pair of jobs to create the similarity matrix used in the genetic algorithm approach. A similarity matrix was generated to evaluate similarity between jobs at each worker level. The matrix generated for the example problem is shown in Table 6.

Table 6. Product similarity coefficients

Products	1	2	3	4	5	6	7	8	9	10
1	-	0.76	0.67	0.5	0.58	0.67	0.76	0.67	0.67	0.76
2	-	-	0.76	0.58	0.76	0.88	0.67	0.67	0.88	0.76
3	-	-	-	0.5	0.58	0.67	0.58	0.58	0.67	0.67
4	-	-	-	-	0.58	0.67	0.58	0.58	0.67	0.5
5	-	-	-	-	-	0.88	0.58	0.58	0.88	0.67
6	-	-	-	-	-	-	0.67	0.67	1.00	0.76
7	-	-	-	-	-	-	-	0.76	0.67	0.58
8	-	-	-	-	-	-	-	-	0.67	0.5
9	-	-	-	-	-	-	-	-	-	0.76
10	-	-	-	-	-	-	-	-	-	-

The capacity requirements are calculated using the production rates found in phase 1 and demand. Each product's demand and capacity requirements are shown in Table 7. Capacity requirements are shown in minutes and as fraction of the total minutes available in each cell. Also, capacity requirements are determined for both the LC and FC and cells are loaded so that neither is over the allowable utilization, 100% for the example problem. An example solution for the problem are shown in Table 8. Cell 1 is assigned products 10, 7, 6, and 2. The utilization in the LC is 0.88 and the utilization for FC is 0.71.

Table 7. Capacity requirements

Product	Demand	Lasting cell			Finishing cell		
		Production rate	Capacity requirement (min)	Capacity requirement (as fraction of 2400 min)	Production rate	Capacity requirement (min)	Capacity requirement (as fraction of 2400 min)
1	1863	2.84	656.70	0.27	4.55	409.86	0.17
2	2147	3.53	608.32	0.25	4.07	528.17	0.22
3	2291	2.07	1107.30	0.46	3.74	612.85	0.26
4	1328	2.51	529.29	0.22	4.04	328.67	0.14
5	2300	2.33	988.99	0.41	5.17	444.67	0.19
6	1627	3.19	509.79	0.21	3.67	443.36	0.18
7	1389	4.07	341.70	0.14	4.55	305.58	0.13
8	2010	3.52	570.84	0.24	2.44	824.11	0.34
9	1601	3.28	488.15	0.20	9.76	164.10	0.07
10	1409	2.07	681.01	0.28	3.17	443.84	0.18

Table 8. Cell loading results for the example problem

	Products	LC utilization	FC utilization
Family 1	[10,7,6,2]	0.88	0.71
Family 2	[8,3,1]	0.97	0.77
Family 3	[9,5,4]	0.83	0.4

5.1 Flowshop Scheduling Methodology

The objective of the mathematical model for flowshop scheduling is to sequence the jobs in each part family such that the desired performance measure (makespan, total tardiness, or number of tardy job) is minimized.

Production rates found in phase 1 are used to determine processing times for the LC and FC. The processing times for the RMC are determined by the shoes mold type and size. The indices, parameters and decision variables are described below.

Index:
u : Cell index;
j : Job index

Parameters:
P_{ij} : Processing time of job j in cell i;
d_j : Due date of job j;
r : Small number j;
M : Large number

Decision variable:
y_{ij} : Start time of job j in cell i;
c_{ij} : Completion time of job j in cell i;
t_j : Tardiness of job j;
w_j : 1 if job j is tardy, 0 otherwise;
z_{ji} : 1 if job j is processed before job i in cell 1, 0 otherwise;
b_{ji} : 1 if job j is processed before job i in cell 2, 0 otherwise;
g_{ji} : 1 if job j is processed before job i in cell 3, 0 otherwise;
TT : total tardiness variable;
h : makespan variable
Objective functions:

$$\min\ N_1 = \sum w_j + r \times \sum t_j, j = 1, 2, \cdots, n \tag{6}$$

$$\min\ N_2 = TT \tag{7}$$

$$\min\ N_3 = H. \tag{8}$$

$$\text{s. t.}\ \sum t_j = TT, j = 1, 2, \cdots, n \tag{9}$$

$$c_{ij} \leq h, j = 1, 2, \cdots, n \tag{10}$$

$$y_{i+1,j} - y_{ij} \geq p_{ij}, j = 1, 2, \cdots, n \tag{11}$$

$$c_{ij} - y_{ij} \geq p_{ij}, j = 1, 2, \cdots, n \tag{12}$$

$$c_{ij} - \geq d_j, j = 1, 2, \cdots, n \tag{13}$$

$$M \times w_j \geq t_j, j = 1, 2, \cdots, n \tag{14}$$

$$y_{ij} - y_{ij+1} + M \times z_{ij} \geq p_{ij}, j = 1, 2, \cdots, n \tag{15}$$

$$y_{ij+1} - y_{ij} - M \times z_{ij} \geq -M + p_{ij}, j = 1, 2, \cdots, n \tag{16}$$

$$y_{ij} - y_{ij+1} + M \times b_{ij} \geq p_{ij}, j = 1, 2, \cdots, n \tag{17}$$

$$y_{ij+1} - y_{ij} - M \times b_{ij} \geq -M + p_{ij}, j = 1, 2, \cdots, n \quad (18)$$

$$y_{ij} - y_{ij+1} + M \times g_{ij} \geq p_{ij}, j = 1, 2, \cdots, n \quad (19)$$

$$y_{ij+1} - y_{ij} - M \times g_{ij} \geq -M + p_{ij}, j = 1, 2, \cdots, n. \quad (20)$$

The various objective functions which can be used depending on the desired performance measure are shown in Eqs. (6), (7), (8), which seek to minimize number of tardy jobs, total tardiness, and makespan, respectively. Eq. (9) sums the jobs tardiness to define total tardiness. Eq. (10) establishes the relationship between the jobs' completion time and its makespan, ensuring that makespan is equal to the completion time of the last job. Eq. (11) ensures that a job has to finish processing in its current cell before it can start in the following cell. In a similar manner, Eq. (12) ensures that a job must complete processing in the final cell before it can be labeled complete. The relationship between completion times, due dates, and tardiness is established in Eq. (13). Equation (14) assigns a value of either zero, for on time jobs, or one, for tardy jobs. Equations (15) and (16) perform a pairwise comparison between jobs in cell 1 to determine which job is processed first. This comparison is done between all jobs and Eq. (17) through Eq. (20) show the comparison being performed in the second and third cells.

The schedules for the part families from the example with regards to makespan are shown in Table 9. For example the sequence for part family 1 is 2-7-6-10 and the makespan is 53. A Gantt chart is shown for part family's schedule with regards to makespan in Fig. 5.

Table 9. Example schedules to minimize makespan

	Schedule				Makespan
Part Family 1	2	7	6	10	53
Part Family 2	8	3	1	-	60
Part Family 3	5	4	9	-	50

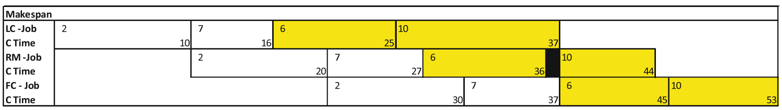


Fig. 5. Gantt chart for family 1

6 Genetic Algorithms Approaches for Cell Loading

A genetic algorithm for cell loading is proposed in this section. Chromosome representation, crossover, and mutation strategies used are discussed. The genetic algorithm has two crossover strategies and the option to perform mutation before crossover. This created four general strategies; (1) order-based crossover before mutation, (2) position-based crossover before mutation, (3) order-based crossover after mutation, and (4) position-based crossover after mutation. Within each of the four general strategies, ten combinations of crossover probability and mutation probability were tested. Five trials were performed for each crossover and mutation probability combination in each of the general strategies.

The chromosome initially is created by randomly ordering the products. The cells are then loaded sequentially from the randomly ordered list. When the utilization in either the lasting cell or finishing exceeds a maximum allowed utilization, that cell is closed and the next cell is opened. This continues until all products are assigned to a cell. An example of the chromosome representation is shown in Table 10. The fitness function is calculated using Eq. (21).

$$\max Z = \sum_{i=1}^{n}\sum_{k=1}^{n} S_{ik}x_{ik} - \sum_{k=1}^{n}(px_{kk}). \tag{21}$$

Table 10. Example chromosome representation

	Product sequence									
	1	2	3	4	5	6	7	8	9	10
Product	2	4	1	5	6	8	3	9	10	7
LC Utilization	0.24	0.3	0.28	0.17	0.22	0.2	0.2	0.15	0.23	0.26
FC Utilization	0.27	0.3	0.17	0.3	0.25	0.22	0.18	0.15	0.23	0.18
Cell	1	1	1	2	2	2	2	3	3	3
Total LC Util	0.24	0.54	0.82	0.17	0.39	0.59	0.79	0.15	0.38	0.64
Total FC Util	0.27	0.57	0.74	0.3	0.55	0.77	0.95	0.15	0.38	0.56

Two crossover strategies are used in the genetic algorithm, order-based crossover and position-based crossover, and these strategies are discussed in the ensuing paragraphs. Position-based crossover is performed by selecting a set of positions at random from one parent chromosome and passing the values at these positions to the child. The missing genes are then selected and filled in the child chromosome in the order they appear in the other parent chromosome. If there are an even number of products, (#Products)/2 is used to determine the number of random positions selected, (#Products-1)/2 is used for an odd number of products. Order-based crossover is performed by selecting a string of genes from one parent and passing the string onto the child chromosome. The remaining genes are then filled in the order they appear in the other parent chromosome. Examples of position-based crossover and order-based crossover are shown in Figs. 6 and 7, respectively. The arrows show how genes are passed from the parent chromosome to the child chromosome.

The mutation strategy used is reciprocal exchange mutation. This is performed by using random numbers to select two positions in the chromosome and swapping the values in these positions. An example of this mutation strategy is shown in Fig. 8.

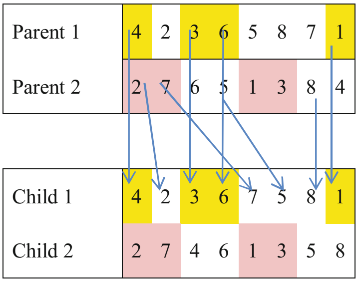


Fig. 6. Position-based Crossover

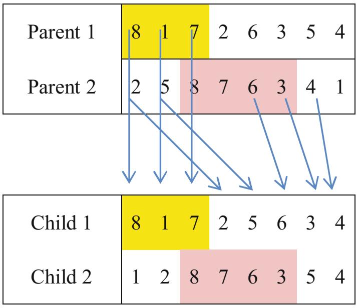


Fig. 7. Order-based Crossover

Chromosome before mutation								
Job	1	2	3	4	5	6	7	8
Chromosome after mutation								
Job	1	7	3	4	5	6	2	8

Fig. 8. Mutation Example

After crossover and mutation have been performed, products are reassigned to cells in the same manner as the original chromosome to ensure the maximum utilizations are not exceeded. Selection is by selecting the top chromosomes among both the parent and child chromosomes to move onto the next generation. This study also investigates the effect of performing mutation before crossover. Classical genetic algorithms flow is crossover before mutation. By reversing the traditional flow, it is possible better parent chromosomes will be created and performing crossover with the better parent chromosomes will result in improved offspring chromosomes. Selection is done in this study by selecting the best ranked chromosomes from all the parent and offspring chromosomes for the next generation.

7 Results and Conclusion

Scheduling is done on the part families found using the four GA strategies. The scheduling results for the worker levels are shown in Tables 11, 12, 13, 14, 15 and 16 below. The bolded and italicized values in the tables below indicate that those are the best result in the particular performance measure in that worker level.

At the 15/20 worker level given in Table 11, GA Strategy 2 either outperformed or tied with respect to MS, TT and number of machines required whereas GA Strategy 4 provided the best result in nT. At the 16/19 worker level shown in Table 12, GA Strategy 4 dominated all other strategies. At the 17/18 worker level presented in Table 13, no strategy dominated others. However, GA Strategy

3 outperformed others with respect to total tardiness and number of machines while GA Strategy 4 outperformed in makespan. At worker level 18/17 given in Table 14, GA Strategy 3 dominated in all scheduling performance measures. However, this strategy requires a significant increase in number of machines relative to the other strategies. This is because an additional part family is opened in this strategy. At worker level 19/16 shown in Table 15, no strategy demonstrated dominance over the rest. GA Strategy 2 and GA Strategy 4 performed equally well. At worker level 20/15 presented in Table 16, GA Strategy 3 dominated in makespan and tied for dominance in nT and number of machines required. Also, GA Strategy 3 provided the lowest makespan among all worker levels in this problem. GA strategy 4 dominated in total tardiness.

The results are summarized in Tables 17 and 18. In terms GA strategies, GA Strategy 3 was the overall best GA Strategy (3 out of 7 times). This was followed by GA Strategies 2 and 4 (2 out of 7 times). GA Strategy 1 never made the list. The variation in MS was 58-61 and nT was 8-9 whereas it was much

Table 11. Worker allocation 15/20 scheduling summary

Solution strategy	15/20 results			
	GA1/MM	GA2/MM	GA3/MM	GA4/MM
MS	60	60	60	60
nT	9	9	10	8
TT	145	137	139	139
# Machines	276	274	274	277

Table 12. Worker allocation 16/19 scheduling summary

Solution strategy	16/19 results			
	GA1/MM	GA2/MM	GA3/MM	GA4/MM
MS	61	61	60	60
nT	9	9	9	8
TT	124	124	123	122
# Machines	285	285	273	272

Table 13. Worker allocation 17/18 scheduling summary

Solution strategy	17/18 results			
	GA1/MM	GA2/MM	GA3/MM	GA4/MM
MS	62	62	62	60
nT	10	9	9	9
TT	122	127	120	126
# Machines	286	277	276	279

Table 14. Worker Allocation 18/17 Scheduling Summary

Solution strategy	18/17 results			
	GA1/MM	GA2/MM	GA3/MM	GA4/MM
MS	66	64	61	63
nT	10	10	9	10
TT	157	147	135	159
# Machines	256	247	275	254

Table 15. Worker Allocation 19/16 Scheduling Summary

Solution strategy	19/16 results			
	GA1/MM	GA2/MM	GA3/MM	GA4/MM
MS	60	59	60	59
nT	9	8	9	8
TT	111	116	111	116
# Machines	279	279	279	279

Table 16. Worker allocation 20/15 scheduling summary

Solution strategy	20/15 results			
	GA1/MM	GA2/MM	GA3/MM	GA4/MM
MS	60	60	58	61
nT	9	9	8	8
TT	128	128	117	116
# Machines	275	275	270	278

Table 17. Summary of manpower level results

	Manpower levels					
	15/20	16/19	17/18	18/17	19/16	20/15
Overall best	GA2	GA4	GA3	GA3	GA2/GA4	GA3
GA strategy						
Best MS	60	60	60	61	59	58
Best nT	8	8	9	9	8	8
Best TT	137	122	120	135	111	116
Best # Machines	274	272	276	247	279	278

Table 18. Summary of performance measure results

Performance measure	Best so far	Manpower levels	GA strategies
MS	58	20/15	GA3
nT	8	20/15, 19/16, 16/19	GA3/GA4, GA2/GA4, GA4
TT	111	19/16	GA1/GA3
# Machines	247	18/17	GA2

higher in terms of TT (111–137) and number of machines (247–279) as shown in Table 17. Finally, the best values of performance measures obtained in this experimentation are summarized in Table 18. We can safely say that unbalanced manpower levels in LC and FC produced better results, in other words, worker levels 20/15, 19/16, and16/19. One can actually say that higher manpower level in LC seemed to work better. One again, GA Strategy 3 produced better results along with GA Strategy 4.

References

1. Akturk MS, Wilson GR (1998) A hierarchical model for the cell loading problem of cellular manufacturing systems. Int J Prod Res 36(7):2005–2023
2. Babayiğit C (2003) Genetic algorithms and mathematical models in manpower allocation and cell loading problem [electronic resource]. Master's thesis, Ohio University
3. Chang OK, Shin HJ (2003) Scheduling jobs on parallel machines: a restricted tabu search approach. Int J Adv Manuf Technol 22(3):278–287
4. Gupta JND, Chantaravarapan S (2008) Single machine group scheduling with family setups to minimize total tardiness. Int J Prod Res 46(6):1707–1722
5. Gupta JND, Ho JC (2001) Minimizing makespan subject to minimum flowtime on two identical parallel machines. Comput Oper Res 28(7):705–717
6. Huang J, Süer GA, Urs SBR (2012) Genetic algorithm for rotary machine scheduling with dependent processing times. J Intell Manuf 23(5):1931–1948
7. Mese E (2009) Cell loading and family scheduling for jobs in a shoe manufacturing company. Master's thesis, Ohio University
8. Nakamura N, Yoshida T, Hitomi K (1978) Group production scheduling for minimum total tardiness part (i). IIE Trans 10(2):157–162
9. Saad SM, Baykasoglu A, Gindy NNZ (2002) A new integrated system for loading and scheduling in cellular manufacturing. Int J Comput Integr Manuf 15(1):37–49
10. Stnha RK, Hollier RH (1984) A review of production control problems in cellular manufacture. Int J Prod Res 22(5):773–789
11. Subramanian AK (2004) Cell loading and scheduling in a shoe manufacturing company. Master's thesis, Ohio University
12. Süer G, Subramanian A, Huang J (2009) Heuristic procedures and mathematical models for cell loading and scheduling in a shoe manufacturing company. Comput Ind Eng 56(2):462–475
13. Süer GA (1996) Optimal operator assignment and cell loading in labor-intensive manufacturing cells. Comput Ind Eng 31(1–2):155–158

14. Süer GA, Saiz M et al (1995) Manufacturing cell loading rules and algorithms for connected cells. Manuf Res Technol 24(24):97–127
15. Süer GA, Cosner J, Patten A (2008) Models for cell loading and product sequencing in labor-intensive cells. Comput Ind Eng 56(1):97–105
16. Süer GA, Ates OK, Mese EM (2014) Cell loading and family scheduling for jobs with individual due dates to minimise maximum tardiness. Int J Prod Res 52(19):5656–5674

Research on the Competitive Strategy of Two Sided Platform Enterprises Based on Hotelling Model

Zeming Wang(✉) and Yinpeng Guo

School of Management, Northwest University for Nationalities,
Lanzhou 730030, People's Republic of China
641881800@qq.com

Abstract. In this paper, to take the group buying platform as an example, we use Helloing model to analyze the influence the pricing and profit of the platform enterprises under the different homing behaviors of users. The analysis platform of enterprise competitive strategy from the perspective of pricing and profit, we show that compared with network externalities, the influence of differences is bigger about platform enterprises; changed the competitive strategy to improve customer loyalty can bring more profits; mutual cooperation between the platform of enterprise will increase the enterprise platform competitiveness. We hope to provide references for the competitive strategic choices of the platform enterprises.

Keywords: Group purchase platform enterprise · Hotelling model · Bilateral platform enterprise · Competitive strategy

1 Introduction

In recent years, bilateral market is the industrial organization and the theory of competitive strategy research. Its unique business platform competition strategies have attracted the attention of scholars. The earliest group buying websites appeared in United States (U.S. buy network Groupon). It is mainly the use of consumers to get a larger discount psychology, and take shape an effective network interaction. Then, group buying to earn service fees from businesses. Influenced by this operating mode Chinese group purchase network development rapid after 2010. From the group purchase network to the mobile terminal APP application, Meituan, Public comment network, Juhuasuan, Baidu Nuomi, Wowo Group and other large professional group buying provide abundant products that involved in almost every aspect of people's life such as food, clothing, cosmetics beauty, fitness, hotel, tourism and so on. Then consumers are attracted by lower prices and larger discounts. With the increase of the number of consumers, more and more merchants registered in platform of enterprise. Volume of business brings in profit for the platform enterprise. According to the Chinese group purchase market statistics in June 2015 shows that by the end of June

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_47

2015, Chinese network group purchase transaction volume reached 16.74 billion yuan, compared to the same period the growth of 182.5%, the number of people who participate in the group reached 250 million person-times, compared to the same period the growth of 161.8%. As a typical enterprise platform, group purchase platform competition is increasingly fierce. Everything is not going smoothly. Due to the diversity of businesses and consumers, some businesses and consumers will choose a platform for business, others will be registered in enterprises with different network platform. The choice behavior of different businesses and consumers will influence the pricing direction and profit of enterprise platform. Research the choice behavior of consumers and businesses to develop a reasonable pricing mechanism and competitive strategy for different types of choice is of great significance for the development of the group buying platform.

2 Related Works

Entering twenty-first Century, the bilateral market theory has attracted wide attention with the rapid development of the information economy. As the leading theory of industrial organization, its importance has been increasing. The earliest examples of two-sided markets can be traced back to the "New York sun" led the so-called "penny movement", all the journalists no longer need to pay the cost to buy a newspaper, the newspaper operator is the main profit from advertising revenue [1]. Rochet and Tirole [7] thought bilateral market is described that have two independent edge, and its ultimate benefit from the platform transaction. They established a bilateral market platform model that has become research basis of bilateral market theory. Cheng and Sun [6] put forward that bilateral market is a dumbbell type structure, the participating parties B expressed as consumers, and the participating parties S as businesses, as shown in Fig. 1.

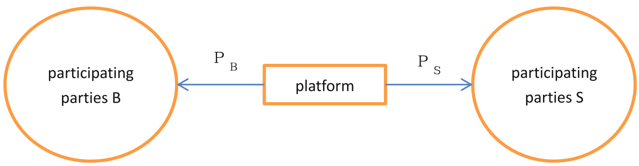


Fig. 1. Two sided market simplified schematic

Chen [4] proposed the network group purchase system, the network group purchase has carried on the more comprehensive description (Fig. 2). As a bilateral platform for enterprises, group buy site is also a two different types of participants through the platform to trade. It is the specific bilateral market dumbbell type structure too.

Through the study a lot of literature, most of scholars think that the bilateral market has the following characteristics. Firstly, the platform has two different types of user (businesses and consumers) at least in the market, and provides

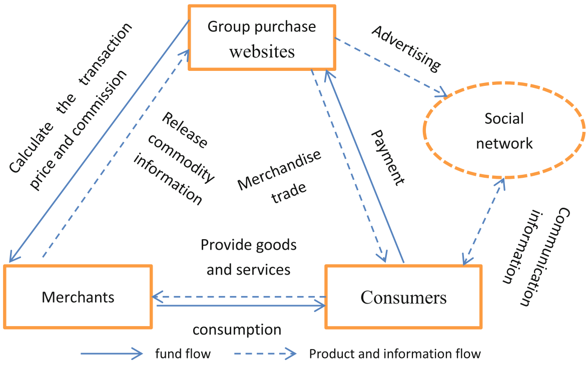


Fig. 2. Network group purchase system

different products or services for different types of user. Or make a deal for different types of user through the platform. Secondly, various types of user interaction platform, they need each other. Platforms help them to complete the transaction and improve the efficiency of transaction. Thirdly, the existence of network externalities between different types of client platform, which increased the number of user on the platform in one side (for example increasing the number of consumers) will increase the value of platform for other types of user [5].

In recent years, with the advent of the era of big data, there has been a new platform of bilateral market enterprise research. Bardey, Helmuth and Lozachmeur [2] studied competition in two-sided markets with a common network externality rather those than with the standard inter-group effects. Behringer and Filistrucchi [3] showed that a two-sided monopolist may find it short-run profit-maximizing to charge a price below marginal cost on one side of the market. Hence showing that the price is below marginal cost on one side of a two-sided market cannot be considered a sign of predation. They then argue for a two-sided Areeda-Turner rule that takes into account price-cost margins on both sides of the market. Roger [8] studies duopoly in which two-sided platforms compete in differentiated products in a two-sided market. Zhu [10] talked about the platforms that is from the one-sided market to the one-sided + two-sided market, as an example of JD.com. Scholars from different angles discussed the bilateral market competition relationship, research more and more comprehensive. This paper will discuss the competitive relationship of bilateral market based on the user ownership relationship.

3 Put Forward Problems

In the bilateral market, users (merchants and consumers) will consider their own interests. Some consumers and businesses only register a group purchase platform, or choose only one platform to make a deal. This user's behavior belongs to the single homing. Other consumers and businesses are registered or select more than one platform to consumption. This user behavior belongs to multi homing. Users behavior present in more than one group purchase platform, because its ownership behavior is not unified. Poolsombat and Vernasca thought in the bilateral market this behavior called partial multi-homing that one part of user belongs to the single homing and other part of user belongs to multi homing. In reality, the users both ends of platform companies such as most of group purchase platform belong to partial multi-homing. Taking into account the factors of customer loyalty, homing behaviors of user are divided into three types: the type of one is the consumers and businesses all belong to single homing behavior; the type of two is the consumers or businesses belonging to partial multi-homing, and other businesses and consumers belong to single homing; types of three consumers and businesses are belonging to partial multi-homing. According to the characteristics of the bilateral market and external factors that influence the bilateral market, this paper studies the following problems.

(1) Owing to the difference between two sides of group platform and impact on network externalities, leading to businesses and consumers have different homing behaviors. The type of users' homing behavior in the difference between the group buying platforms can have an effect on great difference. Whether it can reduce the impact of network externalities, obtain a larger profit for group buying platforms.
(2) Businesses and consumers can freely choose the ownership of the group purchase platform. Belonging to different behavior of businesses and consumers will affect the group purchase platform enterprise pricing. Because of consumers and businesses of different homing types, group purchase platform pricing may be different.
(3) In the two-sided market can choose freely under the condition of different attribution behavior of businesses and consumers. Corporate profits will affect by consumers and businesses of different types of ownership. As a result, group purchase platform gains may be different.
(4) In order to develop itself, group purchase platform may often take different strategy. For example they use market segmentation to attract different consumers and businesses. By taking the difference strategy, ownership behavior of consumer businesses may be changed. Ownership direction guiding of users may bring more profits for the platform of enterprise development.

4 The Model of Competition

Hotelling model is an important analytical model in the theory of industrial organization. Amstrong [9] is an earlier study on the pricing strategy of platform competition in two-sided markets by using the Hotelling analysis framework. In this paper, we also learn from this analysis framework to study competition strategy of the group buying platform.

First of all, we define symbols for the convenience of research and then discuss classification. In this model, service cost of the group buying platform to provide users is ignored.

Notation:

i, j: Two platform enterprises with a competitive relationship;
k, m: Businesses and consumers are located on both sides of the group buying, and obey the uniform distribution on the line k, $m = 1, 2$;
v_k^i, v_k^j: The user register exchange on the group-buying platform i and platform j
p_k^i: Group-buying platform pricing for the users;
t: User to buy platform unit transportation cost (its economic meaning of the difference between the two platforms);
β: Network externalities between the users;
v_0: Users obtain basic utility from platform i or platform j based on the net utility of the platform, if all the users will be registered on the platform to have transaction at least, the v_0 tends to infinity;
n_k^i, n_k^j: The number of single homing user group K on the group buying platform i and platform j;
N_k^i: The number of single homing and the number of multi homing on the group buying platform i of users user group K;
π: The profit of group buying, equal to the number of users of the price multiplied by the price;
ss: Both of two sides are single homing
sm: One side is single homing, another side is partial multi-homing;
mm: Both of two sides are partial multi-homing

4.1 Both of Two Sides are Single Homing

In this model, two group purchase platform enterprise that have competitive relation are supposed, and located at the ends of the line segment [0,1]. Businesses and consumers are users of two sides of group purchase platform enterprise. Through the comparison net income of group buying platform 1 and group buying platform 2, users choose a high net income platform registered and traded. As shown in Fig. 3, making Hotelling model, a user side K to group purchase platform 1 distance is x, and to the group purchase platform 2 distance was $1 - x$. Obviously, the group purchase platform profit equals based utility minus the paid price, minus the cost of transportation, and add users to the group purchase website external effect.

Thus we get v_k^1 and v_k^2:

$$v_k^1 = v_0 - p_k^1 - tx + \beta n_m^1, v_k^2 = v_0 - p_k^2 - t1 - x + \beta n_m^2.$$

A and B is homing zero point on the both of two sides of two group purchase platforms, and in this two points the users obtain same profit. The number of users in the platform $n_k^1 = x$, $n_m^1 + n_m^2 = 1$, so:

$$n_k^1 = 1/2 - 1/2t(p_k^1 - p_k^2) + \beta/2t(2n_m^1 - 1).$$

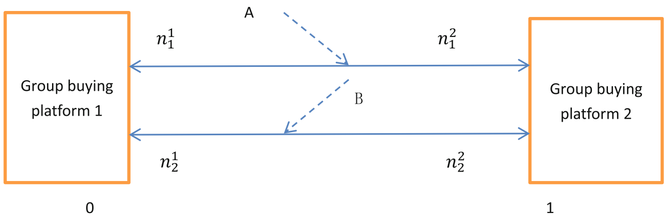


Fig. 3. The user of both of two sides are single homing

Then get the solution:

$$\begin{cases} n_1^1 = \frac{1}{2} - \frac{t(p_1^1 - p_1^2)}{2(t^2-\beta^2)} - \frac{\beta(p_2^1 - p_2^2)}{2(t^2-\beta^2)} \\ n_2^1 = \frac{1}{2} - \frac{t(p_1^1 - p_1^2)}{2(t^2-\beta^2)} - \frac{\beta(p_2^1 - p_2^2)}{2(t^2-\beta^2)}. \end{cases}$$

Because both of two platforms are single homing $n_1^1 + n_2^1 = 1$, $n_1^2 + n_2^2 = 1$, get n_1^2 and n_2^2.

The profit function of the group buying platform is:

$$\begin{cases} \pi_1 = p_1^1 n_1^1 + p_2^1 n_2^1 \\ \pi_2 = p_1^2 n_1^2 + p_2^2 n_2^2. \end{cases}$$

Based on symmetry $p_1^1 = p_1^2 = p_1$, $p_2^1 = p_2^2 = p_2$, p_1, p_2 are got by derivation, as follows, group purchase platform pricing is: $p_1^{ss} = p_2^{ee} = t - \beta$.

Both of sides of the market accounted for 1/2. In the condition of $t > \beta$, profits of the platform 1 and platform 2 have maximum value. So group purchase platform profit is: $\pi^{ss} = t - \beta$

4.2 One Side is Single Homing, Another Side is Partial Multi-homing

Set one side to the side of the single side (Fig. 3 has been described in Fig. 1), part of the multiple homing of 2, the situation as shown in Fig. 4:

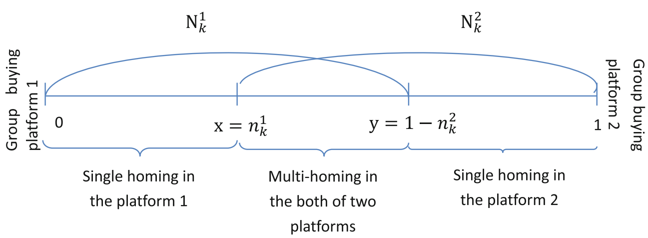


Fig. 4. On both sides of the platform users in the two competing platforms belonging to the range

In the two sides the total number of users is 1. So $N_k^1 + n_k^2 = 1$, $N_k^2 + n_k^1 = 1$. User of single homing on the platform 1 net utility is: $v_k^1 = v_0 - p_k^1 - tx + \beta N_2^1$, K side users of single homing on the platform 2 profit as follows:

$$v_k^1 = v_0 - p_k^1 - tx + \beta N_2^1.$$

The number of users on the platform 1 is: $n_1^1 = \frac{1}{2} - \frac{1}{2t}(p_1^1 - p_1^2) + \frac{\beta}{2t}(N_2^1 - N_2^2)$

Users on the platform 1 of the net utility of n_2^1, which is 2 of the user's distance from the platform 1 of the distance x: $x = n_2^1 = \frac{t-\beta}{t} + (p_2^2)/t + \beta/tn_1^1$.

Users to the platform 2 of the net utility of $1 - n_2^2$, which is 2 of the user's distance from the platform 1 of the distance y: $y = 1 - n_2^2 = \beta/t - (p_2^1)/t - \beta/tn_1^2$.

Simultaneous solution:

$$\begin{cases} n_1^1 = \frac{1}{2} - \frac{t(p_1^1 - p_1^2)}{2(t^2-\beta^2)} - \frac{\beta(p_1^1 - p_1^2)}{2(t^2-\beta^2)} \\ n_2^1 = 1 - \frac{\beta}{2t} - \frac{\beta(p_1^1 - p_1^2)}{2(t^2-\beta^2)} - \frac{\beta^2}{2t(t^2-\beta^2)} p_2^1 - \frac{2t^2-\beta^2}{2t(t^2-\beta^2)} p_2^2 \\ n_2^2 = 1 - \frac{\beta}{2t} - \frac{\beta(p_1^1 - p_1^2)}{2(t^2-\beta^2)} - \frac{\beta^2}{2t(t^2-\beta^2)} p_2^1 - \frac{\beta^2}{2t(t^2-\beta^2)} p_2^2. \end{cases}$$

In the condition of $t > \beta$, the net utility of the group buying platform 1 and group buying platform 2 can get the maximum value, by $p_1^1 = p_1^2 = p_1$, $p_2^1 = p_2^2 = p_2$, group buying platform pricings are: $p_1^{sm} = (t^2 - \beta^2)/t, p_2^{sm} = 0$.

Under the "SM" homing conditions, the numbers of users are:

$$n_1^{1sm} = n_1^{2sm} = \frac{1}{2}, \quad n_2^{1sm} = n_2^{2sm} = 1 - \frac{\beta}{2t}.$$

According to the formula, the profit for the group purchase platform: $\pi^{sm} = \frac{t^2-\beta^2}{2t}$.

4.3 Both of Two Sides are Partial Multi-homing

From the first two sections, we can know the "mm" belonging to the number of users under the conditions of a single:

$$\begin{cases} n_k^1 = 1 + \frac{p_k^2}{t} - \frac{\beta}{t} n_m^2 \\ 1 - n_k^2 = -\frac{p_k^1}{t} + \frac{\beta}{t} n_m^1. \end{cases}$$

Solution to get the user's number of single user on two platforms:

$$\begin{cases} n_1^1 = \frac{t}{t+\beta} - \frac{\beta}{t^2-\beta^2} p_2^1 + \frac{t}{t^2-\beta^2} p_1^2 \\ n_2^1 = \frac{t}{t+\beta} - \frac{\beta}{t^2-\beta^2} p_1^1 + \frac{t}{t^2-\beta^2} p_2^2 \\ n_1^2 = \frac{t}{t+\beta} - \frac{\beta}{t^2-\beta^2} p_1^1 + \frac{\beta}{t^2-\beta^2} p_2^2 \\ n_2^2 = \frac{t}{t+\beta} - \frac{\beta}{t^2-\beta^2} p_2^1 + \frac{\beta}{t^2-\beta^2} p_1^2. \end{cases}$$

Under the circumstances, $p_1^1 = p_1^2 = p_1$, $p_2^1 = p_2^2 = p_2$ we can get the price of the group buying platform:

$$p_1^{mm} = p_2^{mm} = \beta\frac{t-\beta}{2t-\beta}.$$

The number of single belonging to the two group purchase platform is:

$$n_1^{1mm} = n_1^{2mm} = n_2^{1mm} = n_2^{2mm} = \frac{2t^2 - \beta^2}{(t+\beta)(2t-\beta)}.$$

Similarly, in the case of $t > \beta$, the existence of the maximum profit is:

$$\pi^{mm} = \frac{2t\beta^2(t-\beta)}{(t+\beta)(2t-\beta)^2}.$$

5 Analysis of the Model

5.1 The Condition of Maximum Profit

In the condition of $t > \beta$, above of three types of homing are existence of the maximum profit. That is, the difference between the group buying platforms is greater than the network externalities for the maximum value of the platform. These pricings and profits show in Table 1.

Table 1. Three types of users buy site pricing and profits

Type of homing	Pricing	Profits
ss	$p_1^{ss} = p_2^{ss} = t - \beta$	$\pi^{ss} = t - \beta$
sm	$p_1^{sm} = \frac{t^2-\beta^2}{t}$, $p_2^s m = 0$	$\pi^{sm} = \frac{t^2-\beta^2}{2t}$
mm	$p_1^{mm} = p_2^{mm} = \frac{\beta(t-\beta)}{2t-\beta}$	$\pi^{mm} = \frac{2t\beta^2(t-\beta)}{(t+\beta)(2t-\beta)^2}$

The difference of platform enterprises is helpful to eliminate the influence of the uncertainty of network externalities on the profit of the group purchase platform. In other words, if the differences of between the two group buying platforms are enough small, the impact of network externalities may be increased. Then, businesses and consumers are more likely to increase multi-homing behaviors in the platforms, and the uncertainty of platform profit margin is also increased. If there is a big difference between the two platforms of the competition behavior, it can avoid the influence of network externalities on its profitability.

5.2 Different Homing Different Pricing

Compared to the three types of homing structure under the platform pricing: $p_1^{ss} - p_1^{sm} = -\beta\frac{t-\beta}{t}$, due to $t > \beta$, apparently, $p_1^{ss} < p_1^{sm}$. $p_1^{ss} - p_1^{sm} = -\frac{\beta(t-\beta)}{t} < 0$, get $p_1^{mm} < p_1^{ss} < p_1^{sm}$.

$p_1^{mm} < p_1^{ss} < p_1^{sm}$ shows that the bilateral market platform can also by attract consumers by increasing consumer quantity and obtaining consumer loyalty influence of the numbers of registers of businesses at group buying platform. As a result, platform through charge higher fees to businesses to obtain profits.

This conclusion is consistent with the general pricing rules of the group buying platform and the basic characteristics of the bilateral market. There is also difference pricing strategies for the platform enterprises. Adopted differential pricing method is also one of the strategies of competition between enterprises.

5.3 Different Homing Different Profit

Comparison of three types of group purchase websites belonging to profit $\pi^{ss} - \pi^{sm} = \frac{(t-\beta)^2}{2t}$, obviously $\pi^s s > \pi^{sm}$.

$$\pi^{sm} - \pi^{mm} = (t-\beta)\frac{[(t+\beta)(2t-\beta)+2t\beta][(t-\beta)(2t+\beta)]}{2t(t+\beta)(2t-\beta)^2} > 0, \text{ by } \pi^{ss} > \pi^{sm} > \pi^{mm}.$$

Group buying platform can get the highest profit in this condition that users on both of sides belong to single homing. The users on one side belong to single homing, and on another side belong to partial multi-homing take the second place. The behavior of users on both of two sides belong to partial multi-homing is reduced the platform profits. On the one hand, users on the one side belong to single homing and on the another side belong to partial multi-homing can put up the platform enterprise pricing, on the other hand, both sides of the user are single homing can obtain most profitable. This conclusion seems contrary to the conclusion of the 2. In fact, through use of various initiatives to prevent users' multi-homing behaviors, the group purchase platforms enhance loyalty of user for the platform. It is the internal incentive factors for the development of platform enterprises.

5.4 Influence of Difference

Considering the effect of the t for a variety of homing structure and profit pricing platform $\frac{\partial p_1^{sm}}{\partial t} = 1 + \frac{\beta^2}{t^2} > 0 \frac{\partial p_1^{mm}}{\partial t} = \frac{\beta}{(2t-\beta)^2} > 0 \frac{\partial \pi^{sm}}{\partial t} = \frac{1}{2} + \frac{\beta^2}{2t^2} > 0$, $\frac{\partial \pi^{,mm}}{\partial t} = \frac{2\beta^2[t^2(3\beta-2t)+\beta^3]}{(t+\beta)^2(2t-\beta)^3}$, when $t > \beta$, $\frac{\partial \pi^{,mm}}{\partial t}$ number of well over 0, simple that when $\beta \geq \frac{2}{3}t$, $\frac{\partial \pi^{,mm}}{\partial t} > 0$, when $\beta \leq \frac{2}{3}t$, $\frac{\partial \pi^{,mm}}{\partial t}$ well t sign of uncertainty.

From the above of that, users on both of sides belong to single homing and users on one side belong to single homing and another side belongs to partial multi-homing are show that the difference of group buying platform is influence greater than network externalities. In these two cases, pricing and profit of platform will be higher. When users on both of sides belong to partial multi-homing the platform profit is uncertain. This conclusion shows that platform may not be able to get higher profits. If they only consider the unilateral profits and higher prices in the process of competition, they need a reasonable price for platform competition.

6 Conclusion

The users' homing behavior is a complicated economic behavior, which has an important influence on the development of competitive strategy. Based on the group buying platform as an example, from the complexity of social competition in the market, we sum up the characteristic of group purchase platform and discuss the pricing and profit enterprise platform strategy in the different user behavior attribution by making use of the mathematical model. According to the above research conclusions, this paper puts forward some conclusions:

(1) The platform should be based on the needs of users to develop differentiated strategies, personalized competitive strategies to meet the needs of the type of customer needs, and to avoid the impact of network externalities that bring the risk to the platform.
(2) On the online group buying platform Multi homing behavior of users will reduce the fees charged to merchants. Forming malignant price competition and rapid expansion also appeared a lot of homogeneous of group buying platforms such as lashou.com. At last it lost competitiveness. In the course of the operation of the platform, we should pay attention to both businesses and consumers on both sides of the users and try to improve the customers' loyalty for the development of the platform and enterprise profitability.
(3) Users only choose a group buying platform can enhance the competitive advantage of group buying platform. Behaviors of consumers and businesses can get a distinct competitive advantage. Such as the group purchase website, the more good reputations the more consumer users, which can attract more business enterprises to cooperate with the site, forming a virtuous circle and crowding out other competitors.
(4) Effect of multi homing behavior interoperability platform users on the platform can weaken the profit. Cooperation with a number of group purchase website payment platform is not only good for consumer convenience but also attract more business cooperation. Although the consumer is multi homing behavior, in the view of platform, it is a single ownership for the group purchase website. As a whole, the interconnection platform can weaken ownership of consumer behavior, improve the competitiveness of the platform enterprise.

In this paper, we use mathematical model to calculate and compare the competitive price and the profit of the net purchase platform under different ownership types. Because group buying platform is a typical bilateral platform, the study of bilateral platform has a representative significance, to reflect the actual situation of the operation process of the platform. But the deficiency lies in the failure of the general suitability of the model to draw the relevant conclusions. In the future research, we will be in-depth mining related theory, as early as possible to improve the model of universal verification.

References

1. Amery E, Amery M, South LR (2009) American Journalism History: The History of Mass Media Interpretation, 9th edn. Renmin University of China press, Beijing

2. Bardey D, Cremer H, Lozachmeur JM (2014) Competition in two-sided markets with common network externalities. Rev Ind Organ 44(4):327–345
3. Behringer S, Filistrucchi L (2015) Areeda-turner in two-sided markets. Rev Ind Organ 46(3):287–306
4. Chen S (2011) Group buying trends in China. http://wenku.baidu.com/view/81d2ad7ca26925c52cc5bf6a.html
5. Chen Y (2014) Research on the business model of platform enterprises in the bilateral market. PhD thesis, Wuhan University, Wuhan
6. Cheng G, Sun W (2006) The operation mechanism of industry and industry regulation based on bilateral market perspective. Int Financ Res Bank Card 2(1):39–46
7. Rochet JC, Tirole J (1999) Cooperation among competitors: some economics of payment card associations. CEPR Discuss Pap 33(4):549–570
8. Roger G (2016) Two-sided competition with vertical differentiation. J Econ 120: 1–25
9. Shang X, Chen H (2009) Competitive strategy and regulations of platform firms from the perspective of bilateral market. Ind Econ Res 10(6):2461–2472
10. Zhu X, Song B et al (2016) Platforms-from the one-sided market to the one-sided + two-sided market. Bus Trends Digit Era 1:43–61

Research on Tourists Experience in Traditional Culture Festival Activities Based on the Importance-Performance Analysis Research Method

Lu Li[1,2](✉)

[1] Tourism School of Sichuan University, Chengdu 610065, People's Republic of China
342920007@qq.com
[2] School of Economics and Management of Chengdu Technological University, Chengdu 610031, People's Republic of China

Abstract. Chinese excellent traditional culture is an important source of the party in governing the country. Xi Jinping, emphasized the historical influence and significance of Chinese traditional culture several times, and gave it new connotation of the time. Represented by Chengdu Temple Fair in recent years, the traditional culture festival activities were not only for the protection and inheritance of traditional culture, but also became the organic composition part of the local tourism festival activities. In this paper, we stand in the perspective of tourist experience and take Chengdu Temple Fair as the research object. It carried on the questionnaire design and survey to understand of tourists focus and satisfaction. The IPA methods help us to identify the strengths and weaknesses and give improvement suggestions.

Keywords: Tourists experience · Importance-performance analysis (IPA) · Cultural festival activities

1 Introduction

The former researches on traditional culture festival activities mostly focus on culture itself, such as cultural value, cultural inheritance, cultural exchange, and so on. However, in the tourism activities, tourist experience has been considered as one of the most important embodiment of product quality. At the same time, the tourism experience measurement is also the emphasis and difficulty of such research. Therefore, the attention to visitors experience is the crucial factor to make the sustainable development of traditional culture festival activities.

From the perspective of management and marketing, tourist experience is the tourist satisfaction. From the perspective of managers, the tourist experience is formed by a travel destination. At present, tourist satisfaction measurement model in the industry are mainly applied with American consumer

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_48

satisfaction index (ACSI), service quality (SERVQUAL), as well as Importance - performance Analysis method (IPA) [3]. IPA analysis method has been applied to enterprise management and marketing, service quality, competitive analysis, and other fields.

This paper is to study the situation of the festival activities in the tourist experience through the IPA analysis method. It analyses the degree of satisfaction and importance on quadrant distribution. The suggestions for such activities in the future development and quality improvement are also proposed.

Firstly, this paper carries on the literature review of tourism experience and tourist satisfaction, cultural festivals and tourism experience, satisfaction analysis theory to the literature review. It demonstrates the necessity of carry out the research in this field and the feasibility of using the IPA analysis. Then, it focuses in the perspective of tourist experience and takes Chengdu Temple Fair as the research object. It carried on the questionnaire design and survey to understand of tourists focus and satisfaction. Finally, it uses the IPA methods to identify the strengths and weaknesses of traditional culture festival activities and give improvement suggestions.

2 Literature Review

2.1 Tourist Experience and Satisfaction

Research on tourist experience is an important and difficult problem. Tourist experiences have both individual difference and the common performance. The differences mean that the scope, content, depth of each tourists experience are not the same. But if we take the point of tourist groups, their experiences have a certain commonality and show similar regularity.

In theory, the ultimate goal of travel is in order to obtain high quality tourism experience. Therefore, tourism destination and their managers try to achieve the aim through offering high quality tourist experience [2].

From the perspective of management and marketing, tourist experience is the tourist satisfaction. From the perspective of managers, the tourist experience is formed by a travel destination. Thus tourist is the main source of information for destination management. Tourist evaluations of tourism destination constitute the important feedback. Managers take the corresponding management improvement and marketing strategies according to the feedback information [16].

This shows the current research situation for all visitors' experience. Although the ontology of tourists experience has not reached consensus, most scholars formed the understanding of the different Angle of view. They showed their own understandings of tourist experience from different aspects. They also correspondingly gives the measure method and management tools of tourists experience [1,11,15].

2.2 Culture Festival Activities and Tourists Experience

External studies on festival tourism started earlier. Meyer [14] classified tourism festivals according to the festival theme, and made the comparative study of festival tourism between the United States and Canada. Gatz [7], Hoyle [10] studied the problem of festival tourism marketing. Mayfield and Crompton [13] discussed the marketing concept of festival organizers.

It shows that the researches mainly focus on the influence of the large tourism festival, particularly in terms of festival tourism and city relationships [6]. Chinese scholars that studied tourism festival activities focus on the following aspects, including the impact on the destination, the festival tourism operation models, the development present situation and countermeasures, perception of festival tourism for destination residents and tourists, etc.

2.3 IPA Analysis Method

Importance-performance Analysis method, namely IPA method, is widely used for its intuitive, easy to operate and understand. The method is studied out by Martilla and James in 1977 [12].

Since the early 1990s, the IPA analysis began to be widely used in service industry. Its application scope mainly includes service satisfaction, product performance, regional competition (appeal) and macro tourism policy, etc.

On the basis of understanding the multidimensional and multifaceted properties to the satisfaction, IPA analysis method got rapid development in the field of marketing research. It can transmit evaluation message to management through customer evaluation of service product multiple attribute. IPA method asked respondents to evaluate various evaluation factors in all kinds of measures from two aspects of importance and performance.

Fig. 1. IPA matrix

When it is used for satisfaction measurement, performance means the satisfaction evaluation [3]. On the concrete application, first of all, customer understanding the importance of product attributes is to be investigated. Then, we shall understand customer perception of product attribute performance. The overall average of the two will be showed on the two-dimensional coordinate axis [8] (Fig. 1). By the graphic method, managers can accurately grasp the management status, so as to determine the different management area easily.

For the satisfaction evaluation of tourist experience, attribute of tourist products is transmitted into a multilevel index system. The perception of tourism product shows the importance and performance. The status from the survey is put in the I-P model diagram, which can help managers to find the problem and the corresponding management methods to solve the problem. However, the traditional I-P model only presented the direction of management, failure to put forward a measure of the management to improve. Hollenhorst proposed to improvement model-indicator performance estimate (IPE) [9]. IPE model not only to helps managers to understand the importance of each index, but also realize that deviation degree of each index and make its corresponding management action.

3 The Empirical Study

Chengdu Temple Fair is selected as a case for empirical research. Chengdu is the capital of Sichuan province, of Southwest China. It is also one of the most important economic centers, transportation and communication hubs in Western China. Chengdu's earliest temple fair is held before the Qin dynasty period more than two thousand years ago. Since 2005, the Chengdu temple fair festival began to hold in Wuhou Shrine till today. The festival takes public benefit as the core and reveals the local cultural characteristics. Chengdu temple fair festival highlights three kingdoms culture theme, and constantly absorbed all kinds of fresh elements. It gradually formed the culture brand and has a wide influence nowadays. The temple fair was held for years continuously. It has received visitors over tens of millions of people in 11 years.

"The Spring Festival to the Temple fair", it has become a New Year customs of Chengdu area people. Through the successfully held for many years, Chengdu Temple Fair improved year by year. It becomes not only a regional cultural brand, but also one of the most famous Chinese traditional Spring Festival temple fair. Thus, Chengdu temple fair is an appropriate and meaningful case for empirical research. It is both culture and festival tourism activities representative.

3.1 Research Design

The purpose of this study was to apply IPA analysis on festival tourists experience research. Chengdu temple fair festival is selected as empirical research object. The most prominent five factors were extracted in the cultural festivals. Visitors of the target group random questionnaires were distributed during the 2016 Chengdu temple fair. SPASS19.0 software and IPA analysis method was use to analyze tourist experience and satisfaction.

This study used the traditional IPA method to analyze the important factors of tourists' satisfaction at first. Then the importance and satisfaction evaluation correlation detection was carried forward. Thirdly, it made the IPA raw data transformation in the way proposed by Deng [4,5]. Extended importance was calculated by satisfaction evaluation. Each factor was analyzed through transform data. Finally, the two analysis results were compared to get corresponding conclusions and suggestions.

3.2 Survey Design

The questionnaire is made up of three parts. The first part is the survey on tourist satisfaction of cultural festivals elements. On the basis of literature review, the author consulted the expert group of tourism research. According to the specific condition of Chengdu Temple Fair, the tourist experience in cultural festival activities is gained in five elements. They are "The Three Kingdoms culture theme features", "traditional folk performances and activities", "facilities and services", "fashion elements and innovation", and "ticket price setting". Using Likert Scale, we ask the respondents to comment on the five aspects of cognitive performance. In addition, the overall satisfaction perceived evaluation of the elements is added into the survey, so that subsequent calculation and analysis on the importance of extension.

The descriptions of the completely positive statements are made for each element. Respondents are asked to give evaluation ranges from fully agree (5 points) to completely disagree (1 point). So that visitors' satisfaction ranges from fully satisfied (5 points) to completely dissatisfied (1 point). The second part is the survey on the importance of the evaluation factors in cultural festivals activity. In this part, the visitors are asked to the rank the importance by their demands for each element. This design is to reduce the mutual interference of satisfaction and importance evaluation. The third part is the survey of demographic characteristics.

3.3 Data Collection and Analysis

From February to March in 2016, this research made questionnaire investigation at the scene of the Chengdu Temple Fair. A total of five interviewers investigated the visitors in the rest area by random survey. This survey parted out 220 questionnaires, 204 valid questionnaires, and the effective rate of recycling questionnaire was 93%. During the survey, the interviewers required each visitor to score on evaluation factors and overall experience satisfaction. Then, according to the personal value, visitors were asked to make importance ranking for evaluation factors.

After the completion of the questionnaire, the results of the survey were input to SPSS 19.0 database. Data calculation and analysis were done with the help of software. First of all, we made descriptive statistical analysis of the general situation of the respondents. Then, the five key elements of the evaluation of satisfaction and importance of average were calculated. IPA analysis diagram was drawn. The specific steps of research are listed below. The first step, the five elements of satisfaction mean value was determined. Then, according to the importance of evaluation data, we calculated the percentage of the importance, the average value. We made a horizontal axis and vertical axis, the horizontal axis represents the degree of satisfaction axis, the dashed vertical axis represents the importance. The second step, we confirmed the corresponding position location of 5 evaluation factors in four quadrants according to their importance and satisfaction of the actual average. The third step, according to the characteristics of the various factors in different quadrant, we put forward the corresponding countermeasures and suggestions.

In addition, in order to eliminate the correlation between importance and satisfaction factors, many scholars suggested using extended importance to replace self-statement importance. Two kinds of caliber is also analyzed in this paper. Scored for indirect importance, natural logarithmic is computed by partial correlation coefficient of overall satisfaction and satisfaction with all the elements in the way of Wei-jaw Deng, namely extended importance score.

Table 1. Demographic characteristic of tourists

	Item	Number	Percent
Gender	Male	104	51.2
	Female	100	48.8
Age	≤17	6	2.9
	18–30	153	75
	31–45	24	11.8
	46–60	11	5.4
	≥61	10	4.9
Education	Primary school and below	2	2
	Junior high school	14	6.9
	High school	43	21.2
	Undergraduate	138	67.5
	Graduate and above	5	2.5
Monthly income (RMB)	≤3000	21	10.3
	3000–5000	63	30.9
	5000–8000	64	31.4
	8000–10000	24	11.8
	≥10000	29	14.2
Traveler generating region	Chengdu city	83	40.7
	Non - Chengdu city	121	59.3

4 Research Results

4.1 Tourists' Demographic and Tourism Characteristics

Demographic characteristics showed the 204 visitors polled, basically reached the same level of the male to female ratio, which accounted for 51.2% of men, accounted for 48.8% of women. Levels of the age focused on range 18 to 30 years old, accounted for 75%. Level of education was mostly for college/university level, 67.5%. Family income more focused on 3000–8000 interval, accounted for 62.3%. Tourist's origin was in a slightly better non Chengdu city, at 59.3%. Please refer to Table 1 for relevant information.

This study also surveyed the tourists travelling characteristics, including travel patterns, the purpose of the visit, travel expenses, etc. In view of traditional culture festival activities, we also added the play frequency as a survey element. From the statistical results, it can be seen that the first time visitors accounted for 59.6%, but there are still close to half for many times. Chengdu Temple Fair revisit rate is higher, to a certain extent, which also reflects the regional influence of the cultural festival activities. Travel patterns were given priority to with family and friends travel together, accounted for 42.2% and 42.2% respectively. For visit purpose, relax accounted for up to 66.7%, the second is the scenery which is 46.6%. Travel expense (including tickets, transportation, catering and shopping, etc.) average of about 300 Yuan. Relevant information as shown in Table 2.

Table 2. Travel characteristics of tourists

	Item	Number	Percent
Times	1	121	59.6
	2	39	19.2
	3	24	11.8
	>3	20	9.4
Travel patterns	Alone	18	8.8
	With family	86	42.2
	With colleagues or friends	63	30.9
	With lovers	36	17.7
Purpose	Relax	136	66.7
	Family reunion	30	14.7
	Sightseeing	95	46.6
	Photography	9	4.4
	Parents-child campaign	13	6.4
	Friends meeting	23	11.3
	Others	17	8.8
Expense	Average	332.43 (RMB)	

4.2 IPA Analysis of the Importance and Satisfaction

(1) On the basis of tourists' self-statement importance
The results showed that Chengdu temple fair investigated into five elements. Tourists' self-statement importance showed traditional folk performances and activities, Three kingdoms culture theme, fashion elements and innovation, supporting facilities and services, and ticket price setting in turn (See details in Table 3).

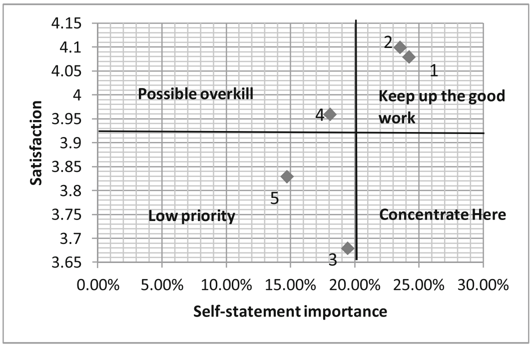


Fig. 2. IPA analysis diagram of self-statement importance and satisfaction

Table 3. Ranking of Tourists' self-statement importance and satisfaction

Ranking of tourists' self-statement importance	Elements (percent)	Importance (5 credits)	Satisfaction	Ranking of satisfaction
1	The Three Kingdoms culture theme features	0.139	4.1	1
2	Traditional folk performances and activities	0.137	4.08	2
3	Fashion elements and innovation	0.115	3.68	5
4	Ticket price setting	0.09	3.83	4
5	Supporting facilities and services	0.087	3.96	3

Thus, we can draw IPA analysis diagram of self-statement importance and satisfaction. It takes tourists self-statement importance as the horizontal axis, satisfaction as the longitudinal axis, and the mean average of each dimension as the midpoint of four quadrants, as shown in Fig. 2. The 2 elements situated in the first quadrant means “Keep up the good work”, namely whose satisfaction and importance are both higher. They are element 1 “Traditional folk performances and activities” and 2 “The Three Kingdoms culture theme features”. The element situated in the second quadrant means “Possible overkill”, namely whose importance is lower but satisfaction is higher. Element The element situated in4 “Supporting facilities and services” is like that. The 2 elements situated in the third quadrant show “Low priority”, namely whose importance and satisfaction are both lower. They are element 3 “Fashion innovation and innovation” and 5 “Ticket price setting”.

(2) On the basis of extended importance
According to the results of the questionnaire, the mean value of overall tourist satisfaction for Chengdu temple fair is 4.00. Then, we calculated extended importance of various factors accordance with the method of Deng [4] (See Table 4).

Table 4. Ranking of extended importance and satisfaction

Ranking of extended importance	Elements	Extended importance (The method of Wei-jaw Deng)	Satisfaction (5 credits)	Ranking of satisfaction
1	The Three Kingdoms culture theme features	0.139	4.1	1
2	Traditional folk performances and activities	0.137	4.08	2
3	Fashion elements and innovation	0.115	3.68	5
4	Ticket price setting	0.09	3.83	4
5	Supporting facilities and services	0.087	3.96	3

As a result, IPA analysis diagram of extended importance and satisfaction was drawn. The basic method is same as the former, but we need to change self-statement importance data into extended importance, as shown in Fig. 3. The 2 elements situated in the first quadrant means “Keep up the good work”, namely still the elements 1 “Traditional folk performances and activities” and 2 “The Three Kingdoms culture theme”. The element situated in the second quadrant is “Possible overkill”, which implied lower importance but higher satisfaction. The elements 5 “Supporting facilities and services” showed that. The element

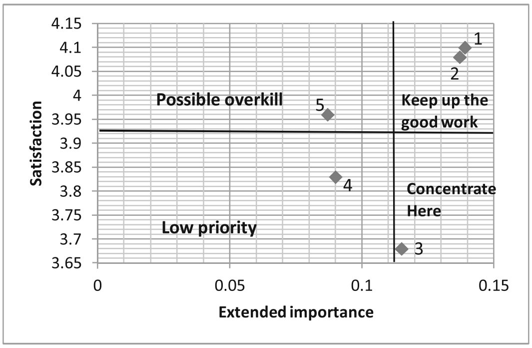


Fig. 3. IPA analysis diagram of extended importance and satisfaction

situated in the third quadrant means "Low priority", which implied extended importance and satisfaction are both lower. It is the elements 5 "Ticket price setting". Element 3 "fashion elements and innovation" is situated in the fourth quadrant, namely lower satisfaction but higher importance, which means "Concentrate Here".

(3) Comparison two kinds of IPA analysis conclusion
The results showed that two kinds of IPA analysis conclusion were basically the same. Slight differences were in two aspects. Firstly, on the ranking of extended importance, "Three kingdoms culture theme" exceed the "Traditional folk performances and activities", in the first place. Extended importance of "Supporting facilities and services" was below "Ticket price setting", which rowed at the bottom. Secondly, by the IPA analysis diagram, we can find that in extended importance analysis, the "Fashion elements and innovation" element was situated in the fourth quadrant, which means "Concentrate Here".

From the two kinds of IPA analysis conclusion, it can be seen that the overall tourist satisfaction of Chengdu temple fair is good. Core competitive elements were traditional folk activities and the Three Kingdoms culture theme features. Tourists' attention and satisfaction on the both elements is high. We need to "Keep up the good work" on the two. And for the evaluation of supporting facilities and services, it was different from traditional scenic spots. Tourists showed lower attention and higher satisfaction. The element was situated in the "Possible overkill" quadrant. The experience length during the festival activities and distance of tourist source region can be related reasons. Ticket price is located in the range of the low awareness and low satisfaction, which embodies the low sensitivity of tickets present price. In addition, the "fashion and innovation elements" is situated in different quadrants of different analysis methods. On the one hand, it reflected the innovation elements had a substantial impact to overall satisfaction. On the other hand, it also indirectly showed the high attention to the traditional culture. Visitors still hope the traditional festival activities can reflect more pure original flavor and essence of traditional culture.

5 Conclusions and Proposal

This study used the IPA analysis to do case investigation and data analysis for tourists experience factors of traditional culture festival activities. We selected Chengdu Temple Fair for empirical research. The self-statement and extended importance were both used to make IPA analysis comparison. Elements need to be focus on and improved were obtained to improve tourist experience evaluation and satisfaction in such cultural festival activities in the future.

First of all, the traditional culture festival activity is a kind of traditional cultural heritage value and means, and also part of the tourism activities. From visitors experience the evaluation results, the majority of visitors still focus on "traditional culture" and "characteristic theme culture" to this kind of festival activities. Chengdu Temple Fair designed a lot of rich and colorful activities, such as opening ceremonies, play god party activities, performing arts activities, theme Lantern Festival, cultural exhibitions, lectures on culture, folk custom activity, interactive games, traditional arts and crafts exhibition, snack sales and other major sections. Visible, the key factor to improve the satisfaction degree of the tourists is to find the core of culture and enhances the core resources of the "traditional culture" in such festival activities.

Secondly, if we take Chengdu Temple Fair as an example, they paid attention to not only the Spring Festival traditional culture in local region, but also its own characteristics and advantages, namely the Three Kingdoms theme culture. In the temple fair, Three kingdoms culture penetrated into all aspects of the festival, from activity forms to content. It has been on full display and extension. Chengdu Temple Fair is different from other temple fairs of the nation's by conducting a lot of activities, such as the Three Kingdoms hero theme day, Three kingdoms culture lamp exhibition, Three kingdoms culture exhibition, the Three kingdoms interactive games, the Three kingdoms cartoon show, the Three kingdoms characters in theme and ambience and so on. The Three kingdoms theme features have been the most prominent, attractive features and core competitiveness of Chengdu temple fair. From the research results, the attention and satisfaction of visitors to the Three Kingdoms culture theme are also one of the best.

Thirdly, the IPA analysis tends to a wide range of applications in the industry due to its simple and easy to interpret. The premise of the analysis request variables on the two dimensions of importance and satisfaction evaluation are independent of each other, and each variable is linear correlation between satisfaction and overall satisfaction. However, in reality survey, the two assumptions are impossible to meet. The main reason is that, on the one hand, the self-statement importance will be inevitably affected by the respondents' satisfaction perception. On the other hand, a certain degree change of satisfaction will not bring the change of overall satisfaction [3]. As a result, this research used two caliber of the self-statement satisfaction and extended satisfaction to do IPA analysis. Managers can make countermeasures for selective in the premise of contrast both the analysis results.

The main limitations of this study is because of selecting importance sorting method for evaluation, we selected limited number of elements from several big aspects. More comprehensive and detailed evaluation and analysis of cultural festivals can be made by kinds of tourist experience factors. These can also become the content of the related research in the future.

Acknowledgements. This study was supported by a grant from Higher School Cultural and Social Science Fund of Sichuan Education Department (to DENG Qingnan) (No. WHCY2016B02).

References

1. Arnould EJ, Price LL (1993) River magic: extraordinary experience and the extended service encounter. J Consum Res 20(1):24–45
2. Bongkoo L, Shafer CS (2002) The dynamic nature of leisure experience: an application of affect control theory. J Leisure Res 34(3):290–310
3. Chen X (2013) The modified importance-performance analysis method and its application in tourist satisfaction research. Tourism Tribune 28(11):59–66
4. Deng WJ, Lee YC (2007) Applying kano model and ipa to identify critical service quality attributes for hot springs hotel in peitou. J Qual 14(1):99–113
5. Deng WJ, Pei W (2007) The impact comparison of likert scale and fuzzy linguistic scale on service quality assessment. Chung Hua J Manag 8(4):19–37
6. Fengying L, Lucang W (2007) Literature review of festival tourism research. Forward Position 7(8):33–35
7. Getz D (1997) Event management & event tourism. Ann Tourism Res 10:22–40
8. Haemoon O (2001) Revisiting importance-performance analysis. Tourism Manag 22(6):617–627
9. Hollenhorst S, Gardner L (1994) The indicator performance estimate approach to determining acceptable wilderness conditions. Environ Manag 18(6):901–906
10. Hoyle LH (2002) Event marketing: how to successfully promote events, festivals, conventions, and expositions. Wiley, New York
11. Manfredo MJ, Driver BL, Brown PJ (1983) A test of concepts inherent in experience based setting management for outdoor recreation areas. J Leisure Res 15(3):263–283
12. Martilla JA, James JC (1977) Importance-performance analysis. J Mark 41(1):77–79
13. Mayfield TL, Crompton JL (1995) The status of the marketing concept among festival organizers. J Travel Res 33(4):14–22
14. Meyer P (1970) Festival: USA and Canada. Ives Washbum, New York
15. Williams DR (1989) Great expectations and the limits to satisfaction: a review of recreation and consumer satisfaction research. In: Outdoor Recreation Benchmark 1988: Proceedings of the National Outdoor Recreation Forum, pp 422–438
16. Yu Xy, Zhu GX, Qiu H (2006) Review on tourist experience and its research methods. Tourism Tribune

How Cross Listing Effects Corporate Performance: Measurement by Propensity Score Matching

Yingkai Tang[1], Huang Huang[1], Hui Wang[1], Yue Liu[2], and Jinhua Xu[3](✉)

[1] Business School, Sichuan University, Chengdu 610064, People's Republic of China
[2] Urban Vocational College of Sichuan, Chengdu 610064, People's Republic of China
[3] School of Management, Guangdong University of Technology, Guangzhou 510520, People's Republic of China
Xujinhua@gdut.edu.cn

Abstract. With the wave of overseas-listed companies returning to A-share market, this paper, taking the perspective of corporate governance, uses Propensity Score Matching (PSM) to measure the difference of operating performance between cross-listed companies and non-cross-listed companies, and uses Structural Equation Model and path analysis to examine the impact mechanism and specific effect of cross listing on operating performance. The results show: (1) Compared with non-cross-listed companies, those companies cross-listed in both A and H-share market, have witnessed a worsening of operating performance, a significant increase of ownership concentration and liabilities to assets (L/A) ratio, and a significant decrease of asset liquidity. (2) Cross listing has affected operating performance mainly through debt structure, asset liquidity, less through ownership structure. The results of this paper can provide the theoretic and empirical basis for future cross listing of companies in China.

Keywords: Cross listing · Propensity Score Matching (PSM) · Structural equation model · Corporate governance · Operating performance

1 Introduction

As the economic globalization accelerates and international capital market becomes more and more integrated, many Chinese companies listed overseas have returned to the domestic A-share market, and got cross-listed. Meanwhile in recent years, domestic and foreign scholars have paid more attention to the change of corporate performance after cross listing. However, empirical results proved cross listing cannot improve corporate performance. Instead, foreign researches showed that after cross listing, corporate internal and external performance had gone down [8,15,26]. Researches on the change of corporate internal

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_49

environment after cross listing, showed cross listing can improve the stock liquidity, reduce the agency cost [6,7], relax financing constraints and reduce the cost of financing constraints [22]. In the past, researches on cross listing often focused on the effect of cross listing on corporate performance, then according to whether the company had been cross listed, set dummy variables for *OLS* regression, ignoring the systematic bias; simultaneously, there was selective bias in samples of cross-listed companies and non-cross-listed companies. No relevant research on impact mechanism of cross listing has been found.

This paper, based on mass samples of China's A-share market, uses Propensity Score Matching (PSM) to examine the difference of operating performance between cross-listed and non-cross-listed companies, and effectively avoid systematic bias in the study of cross listing. And then, this paper uses Structural Equation Model to explore the impact mechanism of cross listing on operating performance.

The structure of this paper is as follows: the first section is an introduction; the second section will be literature review; the third section will be research design; the fourth section will be empirical results; the last section will be conclusion and suggestions.

2 Literature Review

Jaffee and Russell [12] and Stiglitz and Weiss [27] pointed out that there was information asymmetry in credit market, companies were able to complete projects, but banks did not know about this. Most Chinese listed companies get external financing from the credit market, where they encounter serious financing constraints. Therefore, cross listing has always been regarded as an important signal that a company focuses on protecting investors, strictly discloses information and standardizes corporate governance, which can improve public awareness and information environment of the company [2], increase the level of information disclosure in the market, and strengthen investor protection [3,28], thereby reducing the degree of information asymmetry, and relaxing financing constraints. Researches of Lins et al. [18] showed that cross listing can loosen corporate financing constraints, reduce financing costs, and reduce the dependence on internal cash; Doidge et al. [7] thought that compared with non-cross-listed companies, cross-listed companies can expand their financing channels, and prevent management from appropriation of corporate resources; in China, empirical results of Pan and Dai [23] also showed that after cross listing, corporate investments can be much less sensitive to cash flow, so financing constraints can be effectively loosened. Empirical studies of Liu [19] showed that returning to A-share market can ease corporate financing constraints to a certain extent.

With less financing constraints, companies would reduce cash holdings. Gu and Sun [10] pointed out that companies with serious financing constraints had reserve cash to prevent financial risks; Myers [21] pointed out that enterprises holding with a certain amount of cash holdings can seize every investment opportunities, and also can avoid information asymmetry caused by financing constraints. Zhang et al. [30] found that cross listing can enhance the awareness of investor protection, and then reduce the company's cash holdings. However, according to Agency Theory and Free Cash Flow Theory, too much cash holdings can facilitate management to seek personal gain, and increase haphazard investments. La Porta [24] found that with too much cash holdings, management is prone to damage investors' interests by transfer prices, overinvestment, and high salaries. Kalcheva [14] pointed out that too much cash holdings can harm corporate performance and market value.

On the other hand, the mitigation of financing constraints will inevitably lead to the expansion of credit financing. The Pecking Order Theory argues that corporate debt financing outperforms equity financing. Jensen [13] pointed out that debt, as a governance mechanism, can constrain behaviors of major shareholders and management. Berger [4] argued that higher debt ratio can reduce agency costs associated with external equity. However, there are many other factors affecting corporate performance. Aslan and Kumar et al. [1] found that debt financing did not restrain internal agency, but facilitated major shareholders to appropriate corporate resources. The Margaritis and Psillaki [20] pointed out that the relationship of L/A ratio and operating performance presented an inverted U curve, while the ownership concentration was positively correlated to operating performance. Korajczyk and Levy [16] pointed out that without financing constraints, companies would prefer equity financing during periods of economic expansion, and debt financing during periods of economic contraction, and vice versa.

Busaba et al. [5] found when overseas-listed companies return to A-share market in China, they experience poorer post-issuance stock and operating performance in comparison to companies purely listed in China. Meanwhile, companies from less-developed markets would take advantage of the enhanced visibility and prestige associated with their foreign listing to issue shares domestically at inflated prices and favorable terms, and to raise greater proceeds than they can efficiently use. On the basis of Busaba's studies, Kot and Tam [16] pointed out that companies returning to A-share market should be regarded as the valuation and timing of the whole A-share market. Fernandes and Giannetti [9] made sample survey on cross-listed companies worldwide, and drew the same conclusion.

In summary, most scholars held the view about cross listing that after cross listing, operating performance would be damaged, while due to information disclosure, regulatory management and the expansion of financing environment, corporate financing constraints can be mitigated, and corporate governance level can be improved.

3 Research Design

3.1 Research Methodology and Procedures

In order to examine the effect of A+H cross listing, it is necessary to further consider the endogeneity of cross listing. Traditional multiple regression analysis might get biased and inconsistent results; and ordinary matching method cannot solve the problem of endogeneity. This paper uses PSM proposed by Rosenbaum and Rubin [25] to examine the effect of A+H cross listing. With PSM, propensity score of each company is calculated, and the most comparable control samples are matched with samples of A+H cross-listed companies in multiple dimensions; and then the difference between the two groups can show the net effect of A+H cross listing. Therefore, PSM can effectively reduce the selective bias. The specific steps are as follows:

(1) Get propensity scores

Propensity score refers to the conditional probability of a company to get A+H cross-listed.

$$P\left(X\right) = \Pr\left[D = 1\left|X\right.\right] = E\left[D\left|X\right.\right]. \tag{1}$$

D is the research variable, standard samples are divided into AH group and control group, if a company is cross-listed, $D = 1$, otherwise $D = 0$; P is probability of a company to get A+H cross-listed, that is, the propensity score; X is the influencing factor of cross listing, that is, the matching variable.

Then, as what Dehejia did, this paper uses Logit binary regression model to estimate:

$$PS\left(X_i\right) = P\left(X_i\right) = \Pr\left[D_i = 1\left|X_i\right.\right] = \exp\left(\beta X_i\right) / \left(1 + \exp\left(\beta X_i\right)\right). \tag{2}$$

PS is the propensity score of Company i to get cross-listed; D is the research variable, if cross-listed, $D = 1$, otherwise $D = 0$; is the cumulative distribution function of logical distribution; X is the matching variable, which affects cross listing of a company, β is the parametric variable.

(2) Choose the matching method

After getting propensity scores, sample matching is conducted between AH group and control group. Because of the consistency of propensity score, it is difficult to find samples with the same propensity score. Therefore, this paper examines the effect of A+H cross listing by k-nearest neighbors matching.

(3) Conduct balance test

Before getting matching results, it is necessary to test its effectiveness by "hypothesis of commonality" and "hypothesis of independence". "Hypothesis of commonality" requires that control samples should be matched with samples of cross-listed companies by propensity scores. "Hypothesis of independence" requires that there should be no significant difference in the matching variable between AH group and control group.

(4) Calculate the average treatment effect

The average treatment effect (ATT) of AH group and control group, that is, the net effect of A+H cross listing on the research variable, is calculated with the formula Eq. (3). If ATT is statistically significant, then cross listing does have effect on the research variable.

$$\begin{aligned} ATT = & E\left[Y_{1i} - Y_{0i} \left| D_i = 1\right.\right] = E\left\{E\left[Y_{1i} - Y_{0i} \left| D_i = 1, P\left(X_i\right)\right.\right]\right\} \\ & = E\left\{E\left[Y_{1i} \left| D_i = 1, P\left(X_i\right)\right.\right] - E\left[Y_{0i} \left| D_i = 0, P\left(X_i\right) \right| D_i = 1\right]\right\}. \end{aligned} \tag{3}$$

Y is the outcome variable, which is the relevant indicator measuring the effect of A+H cross listing; Y_{1i} and Y_{0i} represent the outcome variable when a company is cross-listed and not cross-listed respectively.

3.2 The Design and Definition of Variables

In case of the measurement bias caused by single indicator, this paper chooses return on total assets (ROA) and Tobin's Q ratio (TBQ) as proxy variables of a company's operating performance, chooses total profitability and market value to measure its operating performance, and ownership concentration (GQJZD), L/A ratio (DEBT) and asset liquidity (XJZCB) to measure corporate governance. Drawing on previous researches, this paper chooses size of asset (SIZE), L/A ratio (DEBT), Tobin's Q ratio (TBQ), fixed asset ratio (GDZCB), debt per share (MGFZ), ownership concentration (GQJZD) and types of enterprises (QYXZ) as matching variables of logistic regression (Logit) (Table 1).

Table 1. The design and calculation of variables

Names of variables	Symbols of variables	Calculation of variables
Return on total assets	ROA	Net return/Average total assets
Tobin's Q ratio	TBQ	Company's market value/Replacement cost of assets
Liquidity ratio	XJZCB	Cash assets at the end of the period/Owner's equity at the end of the period
Size of assets	SIZE	Natural logarithm of total assets
L/A ratio	DEBT	Total debt/Total assets
Fixed asset ratio	GDZCB	Fix assets/Total assets
Debt per share	MGFZ	Total debt at the end of the period/Paid-in capital at the end of the period
Ownership concentration	GQJZD	Sum of squares of the biggest three shareholders' shareholding ratios
Types of enterprises	QYXZ	State-owned enterprises is represented by 1; otherwise, 0

3.3 The Screening of Samples

This paper chooses 58 companies cross-listed in both A and H share market before 2008 as samples of treatment group, and all companies listed in A share market between 2006 and 2014 as initial samples of control group; all samples are screened following three principles: (1) listed financial companies are excluded; (2) companies specially treated and particularly transferred are excluded; (3) companies whose data is missing or abnormal are excluded. Most data of this paper comes from CSMAR Solution, some data comes from corporate annual reports, and the sample size of this research is 14448 companies. All the data is processed with Stata11.2.

4 Empirical Results

4.1 The Screening of Matching Models

In order to make effective matching, regression results of five Logit models are listed in Table 2. Regression results show that the larger size is, and the higher Tobin's Q ratio or the market value is, the more a company is willing to be cross-listed; the higher fixed asset ratio is, the less current asset holdings a company has, then the higher demand for external financing is, therefore, the higher fixed asset ratio is, the more a company is willing to be cross-listed; the higher ownership concentration is, the more a company is willing to be cross-listed, but after cross-listing, the power of controlling shareholders will undoubtedly be diluted, therefore, considering the change of balance of shareholder power after cross listing, we find that cross listing can reduce the power of controlling shareholders, but nearly double the shareholding ratio of the second and third largest shareholders, as a result, ownership concentration index is significantly positive, indicating that major shareholders other than controlling shareholders have great power to decide whether to be cross-listed or not; L/A ratio and debt per share have negative effect on the tendency for cross listing; at the same time, state-owned enterprises (SOEs) are more likely to choose cross listing, as Huang [11] found that, because of China's special capital market environment, leaders of SOEs try to promote their own reputation and benefit their political career through overseas listing; meanwhile, the total asset growth rate has a negative effect on the tendency for cross listing, because controlling shareholders with strong growth ability reject cross listing for fear of the dilution of their power [5].

As what Lian [17] did, this paper, considering the significance of variables in Logit regression, uses Psedo-R2 value and AUC value, the area under the ROC curve, to measure the effect of Logit regression model [29], and all Psedo-R2 values of 5 Logit regression models are more than 0.38, AUC values more than 0.9. When conducting PSM with Logit regression, if AUC value is more than 0.8, then indicators of the regression model are suitable. In this paper, all AUC values are more than 0.9, so indicators are suitable. Considering Psedo-R2 values and the significance of variables, Model 4 is chosen as the matching model.

Table 2. The design and calculation of variables

Variables	Model1	Model2	Model3	Model4	Model5
SIZE	1.2905***	1.3790***	1.2828***	1.3503***	1.3604***
	−26.65	−28.74	−26.57	−26.55	−26.83
DEBT	−1.8016***	−1.4628***	−1.7414***	−1.3864***	−1.4171***
	(−5.16)	(−3.97)	(−5.03)	(−3.74)	(−3.80)
TBQ	0.1169***	0.1230***	0.1221***	0.1212***	0.1138***
	−3.68	−3.31	−3.93	−3.24	−2.92
MGFZ		−0.0153***		−0.0145***	−0.0158***
		(−4.76)		(−4.52)	(−4.87)
CWGG		0.0001			
		−0.02			
ZZCZZL	−1.1183***		−0.9897***		
	(−3.79)		(−3.32)		
GDZCB		0.5561**	0.5441*	0.5156*	0.5156*
		−2	−1.95	−1.85	−1.85
YYSRZZL					−0.0015
					(−0.92)
CWFYL	−0.075				
	(−0.12)				
GQJZD	0.8975**		0.8374**	0.6849*	0.7494*
	−2.23		−2.07	−1.68	−1.86
QYXZ	0.4158***	0.3591***	0.3926***	0.3326***	0.3690***
	−3.6	−3.04	−3.31	−2.79	−3.14
CONSTANT	−31.763***	−34.709***	−32.487***	−34.219***	−34.303***
Psedo-R2	0.3829	0.3858	0.3838	0.3865	0.3858
AUC	0.907	0.902	0.907	0.905	0.905
OBS	14448	14448	14448	14448	14448

Notes: (1) *, **, *** represent significance at the level of 10%, 5% and 1% respectively. (2) The indicator of AUC represents the area under the ROC curve.

4.2 Matching Effect Test

Before getting matching results, balance test was conducted by "hypothesis of commonality" and "hypothesis of independence". Figure 1 is the probability distribution of propensity scores of AH group and control group before and after matching. It shows that before matching, there is a significant difference in the probability distribution of propensity scores between the two groups, and the distribution center of control group is significantly higher than that of AH group. If the difference between the two groups is directly compared, there must be large deviation in the results. However after matching, the probability distribution curve of propensity scores of control group shifts significantly to the right, and the difference of the probability distribution of propensity scores between the two groups significantly decreases, indicating that the matching significantly amends the deviation of the probability distribution of propensity scores, the matching is effective, and hypothesis of commonality is verified.

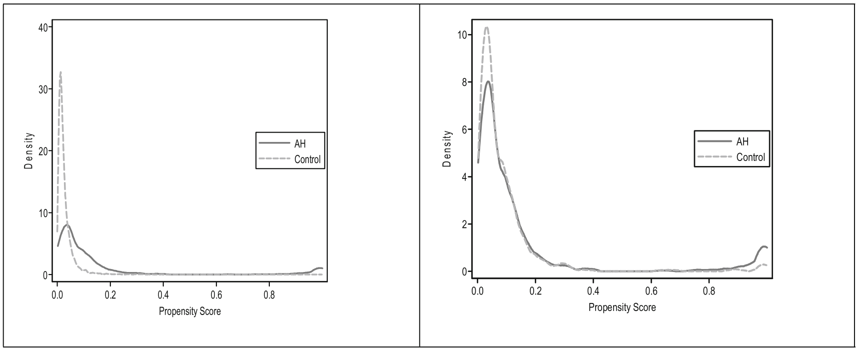


Fig. 1. Probability distribution of propensity scores of *AH* group and control group before and after matching

Hypothesis of independence requires that after matching, there is no significant difference in variables between *AH* group and control group. As shown in Table 3, after matching, all standard deviations sharply decrease, and all absolute standard deviations are less than 10%, indicating that the matching achieves good results. After matching, t-test results show that there is no significant difference between AH group and control group. Above all, hypothesis of balance is verified.

4.3 Univariate Test

After the first phase of PSM, we match cross-listed companies with non-cross-listed companies, and use univariate test to examine their operating performance and corporate governance. In Panel A, all samples in treatment group and control group are compared, and results are as follows: after cross listing in both A and $H - share$ market, the company's cash asset ratio significantly decreases by 0.0406, and ownership concentration significantly increases, indicating that after cross listing, major shareholders increase their stakes in case of the dilution of their power. L/A ratio increases by 0.0398, indicating that after cross listing, it is easier to get external debt financing. Meanwhile, this paper chooses ROA and TBQ as proxy variables of corporate value, which decrease by different degrees after cross listing. In Panel B, samples of state-owned enterprises in treatment group and control group are compared, and results show that cash asset ratio slightly decreases in the significance level of 10%, while none of TBQ,

Table 3. Matching variables and balance test

Names of variables		Mean		Standard deviation (%)	Decrease of standard deviation (%)	*t*-test	
		Treatment group	Control group			*t*	*p*-value
SIZE	Before matching	24.685	21.779	178		49.59	0
	After matching	24.685	24.553	8.1	95.5	1.08	0.28
TQB	Before matching	1.0989	2.055	−57.9		−9.76	0
	After matching	1.0989	1.2167	−7.1	87.7	−1.16	0.763
DEBT	Before matching	0.5854	0.46009	61		12.42	0
	After matching	0.5854	0.5751	5	91.7	0.82	0.413
GQJZD	Before matching	0.2895	0.16866	94.1		20.67	0
	After matching	0.2895	0.28644	2.4	97.5	0.3	0.763
GDZCB	Before matching	0.2976	0.24184	28		6.68	0
	After matching	0.2976	0.31491	−8.7	69	−1.15	0.25
MGFZ	Before matching	14.522	4.8996	39.7		20.79	0
	After matching	14.522	14.263	1.1	97.3	0.13	0.899
QYXZ	Before matching	0.4087	0.24021	36.6		8.28	0
	After matching	0.4087	0.41087	−0.5	98.7	−0.07	0.947

Note: *, **, *** represent significance at the level of 10%, 5% and 1% respectively.

Table 4. The effect of cross listing in *A* and H-share market

Variables	ROA	TBQ	XJZCB	GQJZD	LEV
Panel A Treatment group VS Control group					
Treatment group (AH = 1)	0.0371	1.0988	0.1098	0.2894	0.5854
Control group (AH = 0)	0.0492	1.3462	0.1504	0.254	0.5456
Treatment group-control group	0.0121	0.2474	0.0406	−0.0354	−0.0398
T-value	3.0436***	2.2772**	5.2667***	−3.2978***	−2.6576***
Panel B state-owned enterprises: treatment group vs control group					
Treatment group (AH = 1)	0.0407	1.22	0.1132	0.2972	0.5922
Control group (AH = 0)	0.0446	1.3158	0.1348	0.3106	0.5729
Treatment group-control group	0.0039	0.0958	0.0216	0.0134	−0.0193
T-value	0.6567	0.4833	1.8230*	0.8631	−0.8669
Panel B Private Enterprises: Treatment group VS Control group					
Treatment group (AH = 1)	0.0346	1.0151	0.1074	0.2841	0.5807
Control group (AH = 0)	0.0526	1.3692	0.1622	0.2112	0.525
Treatment group-control group	0.018	0.3541	0.0548	−0.0729	−0.0557
T-value	3.4034***	2.9402***	5.4136***	−5.1239**	− 2.7719***

Note: *, **, *** represent significance at the level of 10%, 5% and 1% respectively, and T-value represents the statistics of *t* in the univariate test.

ROA, ownership concentration and L/A ratio passes the significance test. In Panel C, samples of private enterprises in treatment group and control group are compared, the change trend of each indicator is in accordance with that in Panel A, and all indicators are significant at the level of 1%. However, after cross listing, ownership concentration and L/A ratio of private enterprises are apparently higher than that in Panel A, cash asset ratio decreases more than that in Panel A, ROA and TBQ also decrease more than that in Panel A, indicating that after cross listing, major shareholders of private enterprises have stronger awareness of the risk of ownership dispersion, are more willing to increase their stakes, and are more prone to expand the ratio of debt financing, and to use financial leverage to gain profits (Table 4).

4.4 Descriptive Statistics

Table 5 is the descriptive statistical analysis of samples after matching. The mean of market values is 39 billion yuan ($e^{2}4.3864$), indicating that these companies are big in size, which reflects the choice of overseas-listed companies; their L/A ratios are relatively higher, and the mean of L/A ratio is 56.94%; the mean of return on total assets is 0.0419, and TBQ, 1.1603, which reflect corporate operating performance; the mean of sum of squares of the biggest three shareholders' shareholding ratios is 0.2752, indicating that the ownership concentration of China's listed companies is relatively higher, which reflects the possibility of the practice of what one person says counts. Relevant variables are winsorized.

Table 5. Descriptive statistics

Variables	Mean	Standard deviation	Minimum	P25	P50	P75	Maximum
ROA	0.0419	0.0542	−0.2969	0.0126	0.0331	0.067	0.3113
TBQ	1.1603	1.1849	0.0541	0.4165	0.7628	1.4607	6.5051
XJZCB	0.1261	0.1067	0.0029	0.0506	0.0979	0.1662	0.6283
SIZE	24.3867	1.8899	19.6146	23.1212	24.4368	25.4237	30.6755
DEBT	0.5694	0.2045	0.0581	0.423	0.575	0.7307	0.9642
GDZCB	0.2902	0.216	0.0011	0.1033	0.2323	0.4801	0.8621
MGFZ	11.5098	15.0831	0.514	2.8933	6.4708	12.9131	77.8973
GQJZD	0.2752	0.147	0.0057	0.1669	0.2659	0.3746	0.7994
QYXZ	0.4174	0.4934	0	0	0	1	1

4.5 Path Analysis of Operating Performance

(1) The establishment of conceptual model of cross listing's effect on operating performance

Above all, cross listing will worsen operating performance. This paper will explore the impact mechanism of crossing listing on operating performance from debt structure, asset liquidity and ownership structure.

- Debt Structure. It is certain that changes of financing environment can affect a company's asset structure, and then operating performance.
- Asset Liquidity. The mitigation of financing constraints after cross listing can improve the company's cash holdings, and then affect operating performance.
- Ownership Structure. After cross listing, financing market is larger, ownership structure changes, and corporate governance is improved, and then operating performance is affected.

(2) Path effect analysis

Using Structural Equation Model to perform an iterative scheme for maximum likelihood with the conceptual model of Fig. 2, and path analysis results are shown in Fig. 3. The probability of the whole model in adaptability test, is $P = 0.902 > 0.1$, which does not pass the significance test and accept the null hypothesis, indicating that the theoretical model and sample data are adaptable. Other adaptability indicators are RMSEA $= 0.000 < 0.05$, CFI $= 1.000$, TLI $= 1.025$, IFI $= 1.002$, RFI $= 1.000$, NFI $= 1.000$, indicating that adaptability of the model is great.

① The Effect of Cross Listing on Operating Performance

As shown in Fig. 3, cross listing has no significant direct effect on proxy variables of operating performance ROA and TBQ. Results of significant path effect are listed in Table 6.

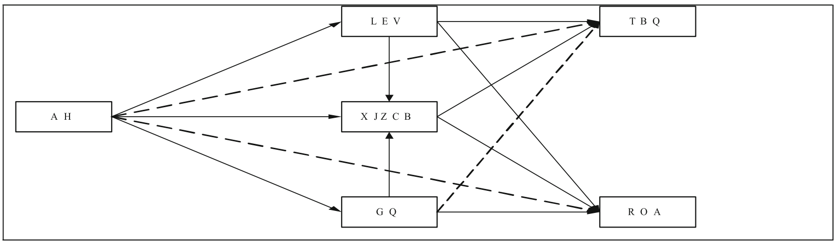


Fig. 2. Conceptual model of cross listing's effect on operating performance

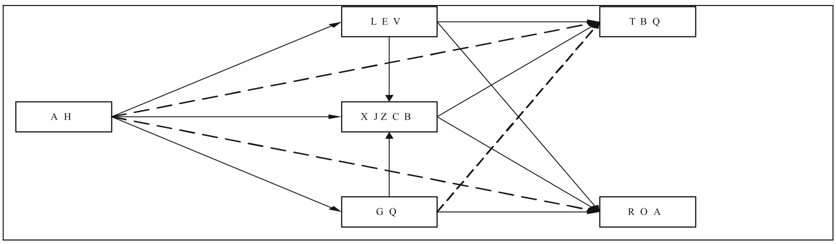


Fig. 3. Coefficient map of path analysis

Table 6. The effect of cross listing on corporate value

Types of path	Paths	Results
Path of ROA	ah-lev-roa	$0.040 \times (-0.111) = -0.0044$
	ah-xjzcb-roa	$-0.031 \times 0.108 = -0.0033$
	ah-gq-roa	$0.035 \times 0.026 = 0.0009$
	ah-lev-xjzcb-roa	$0.040 \times (-0.120) \times 0.108 = -0.0005$
	ah-gq-xjzcb-roa	$0.035 \times (-0.128) \times 0.108 = -0.0005$
	Total effect	-0.0078
	ah-lev-tbq	$0.040 \times (-3.122) = -0.1249$
	ah-xjzcb-tbq	$-0.031 \times 2.308 = -0.0632$
Path of TBQ	ah-lev-xjzcb-tbq	$0.040 \times (-0.120) \times 2.038 = -0.0098$
	ah-gq-xjzcb-tbq	$0.035 \times (-0.128) \times 2.038 = -0.0091$
	Total effect	-0.207

As shown in Table 6, the total effect of path of ROA is -0.0078, indicating that after cross listing in both A and H-share market, corporate operating performance turns down, specifically, after cross listing, debt structure and asset liquidity change in different degrees, affecting operating performance. The total effect of path of TBQ is -0.2070, indicating that after cross listing, corporate market value is negatively affected, and decreases sharply. The coefficient of ah-lev is 0.040, indicating after cross listing, because of the relaxation of financing environment, companies are more inclined to increase debt financing to get rapid expansion; the coefficient of ah-xjzcb is negative, indicating after cross listing, companies reduce their cash holdings to lower the agent cost; the coefficient of ah-gq is significantly positive, indicating after cross listing, ownership concentration increases, as market environment is more complex, and major shareholders are more willing to increase their stakes to further consolidate their power.

As shown in ah-lev-roa and ah-lev-tbq, cross listing can significantly improve debt capacity, indicating cross listing can enhance companies' ability to obtain external financing, however, the increase of L/A ratio cannot bring about growth, but become an obstacle to the development. As shown in ah-gq-roa, after cross listing ownership concentration increases, which improves corporate value, but with limited effect. As shown in ah-xjzcb-roa and ah-xjzcb-tbq, after cross listing, cash asset ratio sharply decreases, and low cash asset ratio has negative effect on corporate value. Simultaneously, as shown in indirect path, after cross listing, ownership concentration and L/A ratio also lead to the significant decrease of cash holdings, and then the decrease of corporate value.

② Path Decomposition of the Effect of Cross Listing on Corporate Value

Empirical researches on A+H cross-listed companies in China, choosing ROA and TBQ as proxy variables of corporate value, show that after cross listing, internal assets structure is not improved. Specifically, when ROA is the proxy variable of corporate value, path of debt contributes 51.6% to the decrease of corporate value after cross listing, path of cash flow, 34.0%, and path of ownership, 14.4%. When TBQ is the proxy variable of corporate value, path of debt structure contributes 65% to the decrease of corporate value, path of cash flow, 30.5%, and path of ownership, 4.5%, which is relatively lower (Table 7).

Table 7. Path decomposition of the effect of cross listing on corporate value

Proxy variables	Types of path	Paths	Effect	Contribution degree	Ratio of contribution degree
Path of ROA	Path of debt	ah-lev-roa	−0.0044	0.0045	46.40%
		ah-lev-xjzcb-roa	−0.0005	0.0005	5.20%
		Total	−0.0049	0.005	51.60%
		ah-gq-roa	0.0009	0.0009	9.20%
	Path of ownership	ah-gq-xjzcb-roa	−0.0005	0.0005	5.20%
		Total	0.0004	0.0014	14.40%
	Path of cash flow	ah-xjzcb-roa	−0.0033	0.0033	34.00%
	Total effect		−0.0078	0.0097	100%
Path of TBQ	Path of debt	ah-lev-tbq	−0.1249	0.1249	60.30%
		ah-lev-xjzcb-tbq	−0.0098	0.0098	4.70%
		Total	−0.1349	0.1349	65%
	Path of ownership	ah-gq-xjzcb-tbq	−0.0091	0.0091	4.50%
	Path of cash flow	ah-xjzcb-tbq	−0.0632	0.0632	30.50%
	Total effect		−0.207	0.207	100%

Note: Contribution degree is indicated by absolute effect, without differentiating positive effect and negative effect.

5 Conclusion

This paper discusses the development status of companies cross-listed in both A and H-share market in China, and examines the effect of cross listing on operating performance by using Propensity Score Matching (PSM), and then uses Structural Equation Model and path analysis to explore the impact mechanism and specific effect of cross listing on operating performance from paths including debt structure, asset liquidity and ownership structure. The results show: (1) Compared with non-cross-listed companies, those companies cross-listed in both *A* and H-share market, have witnessed a worsening of operating performance, a significant increase of ownership concentration and liabilities to assets (L/A) ratio, and a significant decrease of asset liquidity. (2) Cross listing has affected operating performance mainly through debt structure, asset liquidity, less through ownership structure. Main contributions of this paper are as follows: (1) using Propensity Score Matching (PSM) to solve the problem of endogeneity in the research of cross listing; (2) further analyzing the impact mechanism and specific effect of cross listing on corporate performance.

Suggestions are as follows: Firstly, A+H cross listing cannot improve operating performance of Chinese companies, because most A+H cross-listed companies choose to be listed in H-share market first, and then in A-share market, however, compared with the capital market in Hong Kong, the capital market in the mainland of China is not completed enough, thus after A+H cross listing, companies' governance environment has not been improved. Therefore, the top priority is to improve the capital market system. Secondly, domestic listed companies should be encouraged to go out into the developed capital market for financing, utilizing good markets in developed countries, so as to improve the domestic market in an indirect way; after cross listing, companies' ownership concentration and ratio of debt financing significantly increase, indicating debt financing becomes more important, and then companies face greater investment risk. As path of debt is the main factor worsening operating performance, so companies should adjust their ratio of debt financing to a reasonable range, in order to avoid excessive leverage risk; after cross listing, with less financial constraints, companies tend to reduce their cash holdings, and then to make blind investments. As decrease of asset liquidity is also a main factor reducing corporate profitability, companies should properly hold liquidity assets in case of need.

Acknowledgements. The paper is funded by Ministry of Education (AACZXX-01) and Sichuan University (SKGT201602, HZRK201402).

References

1. Aslan H, Kumar P (2016) Controlling shareholders, ownership structure and bank loans
2. Baker H, Nofsinger J, Weaver D (2002) International cross-listing and visibility. J Finan Quant Anal 37:495–521 (in Chinese)
3. Barzuza M (2006) Lemon signaling in cross-listing
4. Berger A, Di Patti E (2013) Computation of discharge through side sluice gate using gene-expression programming. Irrig Drain 62:115–119 (in Chinese)
5. Busaba W, Guo L et al (2015) The dark side of cross-listing: a new perspective on cross-listing. J Bank Finan 57:1–16 (in Chinese)
6. Coffee J (2002) Racing towards the top? The impact of cross-listings and stock market competition on international corporate governance. Columbia Law Rev 102:1757–1831 (in Chinese)
7. Doidge C, Karolyi G, Stulz R (2004) Why are foreign firms that are listed in the U.S. worth more. J Finan Econ 71:205–238 (in Chinese)
8. Durand R, Gunawan F, Tarca A (2006) Does cross-listing signal quality. J Contemp Acc Econ 2:48–67 (in Chinese)
9. Fernandes N (2014) On the fortunes of stock exchanges and their reversals: evidence from foreign listings. J Finan Intermediation 23:157–176 (in Chinese)
10. Gu N, Sun J (2009) Financing constraints, cash flow volatility and corporate precautionary cash holdings. J Bus Econ 4:73–81 (in Chinese)
11. Hung M, Wong T, Zhang T (2012) Political considerations in the decision of chinese soe to list in Hong Kong. J Acc Econ 53:435–449 (in Chinese)

12. Jaffee D, Russell T (1976) Imperfect information, uncertainty, and credit rationing. Q J Econ 90:651–666 (in Chinese)
13. Jensen M (1986) Agency costs of free-cash-flow, corporate finance, and takeovers. Am Econ Rev 76:323–329 (in Chinese)
14. Kalcheva I, Lins K (2007) International evidence on cash holdings and expected managerial agency problems. Rev Finan Stud 20:1087–1112 (in Chinese)
15. King M, Segal D (2005) Are there longer horizon benefits to cross-listing untangling the effects of investor recognition, trading and ownership
16. Kot H, Tam L (2016) Are stock price more informative after dual-listing in emerging markets evidence from Hong Kong-listed chinese companies. Pacific-Basin Finan J 36:31–45 (in Chinese)
17. Lian Y, Su Z, Gu Y (2011) Evaluating the effects of equity incentives using PSM: evidence from china. Front Bus Res China 5:266–290 (in Chinese)
18. Lins K, Strickland D, Zenner M (2005) Do non-US firms issue equity on us stock exchanges to relax capital constraints? J Finan Quant Anal 40:109–133 (in Chinese)
19. Liu X, Tian M, Zhang C (2016) Has returning to a-share market eased the corporate financing constraints? Based on the analysis of cash-cash flow sensitivity
20. Margaritis D, Psillaki M (2010) Capital structure, equity ownership and firm performance. J Bank Finan 34:621–632 (in Chinese)
21. Myers S, Majluf N (1984) Corporate financing and investment decisions when firms have information that investors do not have. J Finan Econ 13:187–221 (in Chinese)
22. Pagano M, Roel A, Zechner J (2002) The geography of equity listing: why do companies list abroad? J Finan 57:2651–2694 (in Chinese)
23. Pan Y, Dai Y (2008) Dual listing and financing constraints-an empirical evidence based on chinese companies with ‘a+h’ dual-listing. China Ind Econ 5:139–149 (in Chinese)
24. Porta R, Lopez-de Silanes F et al (2002) Investor protection and corporate valuation. J Finan 57:1147–1170 (in Chinese)
25. Rosenbaum P, Rubin D (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70:41–55 (in Chinese)
26. Sarkissian S, Schill M (2009) Are there permanent valuation gains to overseas listing. Rev Finan Stud 22:371–412 (in Chinese)
27. Stiglitz J, Weiss A (1981) Credit rationing and markets with imperfect information. Am Econ Rev 71:393–410 (in Chinese)
28. Wojcik D, Clark G, Bauer R (2004) Corporate governance and cross-listing: evidence from European companies
29. Zhang H, Liu J (2015) Does high tendency for cash holdings hurt corporate value of private enterprises? A study based on financing constraint hypothesis or free cash flow hypothesis. Macroeconomics 5:100–109 (in Chinese)
30. Zhang J, Cheng Z, Zhang J (2011) The impact of a cross-listing on the level and the value of cash holdings-evidence from Chinese listed companies. J Shanxi Finan Econ Univ 11:108–115 (in Chinese)

Factors Affecting Employee Motivation Towards Employee Performance: A Study on Banking Industry of Pakistan

Abdullah Khan[1], Shariq Ahmed[2], Sameer Paul[3], and Syed Hasnain Alam Kazmi[3,4](✉)

[1] Khadim Ali Shah Bukhari Institute of Technology, Karachi, Islamic Republic of Pakistan
abdullah@kasbit.edu.pk
[2] Meezan Bank Ltd., Karachi, Islamic Republic of Pakistan
[3] Mohammad Ali Jinnah University, Karachi, Islamic Republic of Pakistan
[4] Southwest Jiaotong University, Chengdu, China
hasnain.alam@jinnah.edu

Abstract. The study provides the cognitive support for developing employee's motivation level at optimized level through consciousness towards applied approach and unrealistic ideas. The diligence of the research is to present the factors that help to motivate the employees to perform their job well. Whereas, the study restrains empirical and descriptive analysis of variables which comprehensibly highlight the significant relationship among the independent variable and dependent variable; the close-ended questionnaire is constructed to compute the statistical data. The study identifies the strong relationship among the defined factors that enhance the employee's motivation level in banking sector. Although the research is also aware the management and the managers of banks about the most preference factors that improves the employee motivation in the banks. This study valued the integrating knowledge, experience, action to groom skilled employees in the banks. Moreover, the main contribution of the research evaluating the role of the motivational factors in banking industry.

Keywords: Motivation · Empowerment · Job environment · Banking · Employee performance algorithm

1 Introduction

In the most current scenario employees have became the primary strength of any business where the employees deliver their continuous effort to drive organizations' decisions into action to accomplish its objectives. Whereas, the motivational factor regarding the employees is becoming a part of organizational strategy. However, motivation is an aspect that helps an individual to select or deselect the job, to continue and work proficiently during his/her job [16].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_50

Further, the motivational research proposes to facilitate the individuals to get motivated through the target to satisfy their unconvinced desires; consequently they apply their endeavor hoping to be satisfied with their needs. Conversely, modern studies' regarding the motivation doesn't highlight the Basel desire conception, but in addition they highlight the objective, equality, and values [24]. However, the after effect from the motivated workforce provides the less employees, turnover, retention and high productivity rate, which leads the organization towards growth and development [13].

Further, the motivation explains as a dynamic strength among individual which inspire the employees to do the incredible work to achieve the goal that regulates to accomplish want or anticipation [19]. It is a complex subject of human behavior which differs from person to person; therefore, varied individuals will be stimulated in unique ways [12]. Individual's behavior will state that what encourages them to implement the effect of both capacity and inspiration [19]. Spiral inspiration can be characteristic and inappropriate [9]. A few theories of inspiration were created to recognize the variables of inspiration that impact human conduct in distinctive ways.

While in conventional management approach the idea is that "Manager recognize the job optimum" is prevailing in current management. In modern conception the representative's inspiration will be extremely enormous particularly for service businesses. While, the resource provokes the representatives, that analyst separated them into two gatherings; outside and inside components [10]. However, external elements, there will be inspiration implies such as meeting expectations condition, pursuing, organization's picture, work assurance, advancement, social environment, and position. In terms of learning organizations, when seeing outer ecological components quick alteration and difficulty to be measured in a firm national and global struggle, these organizations taking choices to profound impact in organizations' by keeping up its presence.

The method individual responds towards organizational working condition depends over job reliability, job stress, working hours. Thus interpersonal associations imitate satisfaction in correspondence with a subordinate. Job satisfaction is completely prejudiced by range of professional characteristic such as organizational ethnicity, interpersonal affairs, worker empowerment, cooperation, and self-apparent accountability [11]. Further, employees assess their working ability by receiving rewards furthermore identify additional fulfillment by getting fair pay in exchange of their efforts within the organization [8]. Fair payment for their efforts can effect in motivation and fulfillment for invulnerable workforce. An individual value system is being studied while creating the awareness of rewards and compensation, whereas individual's values of receiving pride for their efforts to support and increase their work level while it's related to their educational period or from salaried time line [7]. Individual at its initial level of its career pays high values to intrinsic and extrinsic feature related to their job. Moreover, individuals with the startup of their careers valued the rewarding towards their work assessment and environment further the expert and experienced individuals with high potential and opportunities also provide values. As the different researcher consider the motivation as the function of compensation and belief towards their jobs; because of which the high rewarding employees having the high satisfaction in respect to their job [11].

2 Literature Review

2.1 Motivation

Motivation is mostly differentiate as one of the most influential predictors of individual behavior and a key analyst of performance for essential aspect of behavior [22]. Hence, it is not shocking that motivation show in a range of regulation paper [15], scholars and executives have vast attention in accepting person motivation to use communal medium and tools on behalf of companies [14]. Motivational aspect (such as success, appreciation, accountability, effort, progression, individual development) are associated to the application contented though hygiene aspects (such as corporation strategy and management, interpersonal relatives, working environment, pay, job safety, position, reimbursement) are associated to the employment background [6]. Herzberg measured hygiene aspects as extrinsic features, which have an unsettling outcome on the staff job approach and create them eventually displeased in their profession when the needs aren't sufficiently meet. Further, motivational aspects are recognized as intrinsic feature, which make staff fulfilled when needs are meet, conversely, do not formulate them disgruntled in the deficiency of these [21]. Further the study affirms that motivation exists merely when individual appreciates an optimistic correlation that endeavor direct to work performance and work performance directs to incentives [6]. Moreover, the objective locale theory of motivation, where the theory emerge from the proposal of anticipation theory as a target locale is an essential instrument, which work as an "instant valve of individual act" [17] that direct the individuals towards attaining the objective. The objective influences the performance by straight concentration, organizing endeavor, rising determination and inspiring plan. Further the earlier research [3] offer a perceptive of an individual attitude, prospect, plan and aims in life. Moreover, this research helps business to considerate their employees' objective in both their qualified and private life, whilst identifying the need to stimulate them. At the Similar time, they seem to improve the participation, reliability, and assurance of the staff to their businesses. Positively, accepting the features that influence the motivation of staff is an huge dispute for business, leaders and executives [23]. Especially the people worth different aims in life. People might have diverse needs and wants, ethics, ambition, objective, and prospect in life [1]. A motivational presumption effort to clarify the aspect that have straight or circuitous pressure on motivation and business performance, such as worker stimulus and other motivational factors [23].

2.2 Job Environment

The environment where the job performs in their work surroundings can be diverse from those fully contented challenging and unsafe for the life and health of employees. Whereas, the complex operational surroundings can be prejudiced by:

- Peripheral aspect that comprises environment, meteorological condition, temperature, clamminess, breeze, illumination in the office and intrusion, gases and other harmful factors.
- Individual aspect that contain sexual characteristics and maturity of the employee, exhaustion, repetitiveness, adverse stance throughout exertion.
- Another aspect associated with the business of production towards the period of the working hour, working plan, working occasion, work speed, extreme injure, etc. further, the work with complicated operational circumstances may achieve only the workforce who convene precise requirements in stipulations of maturity, gender, education, strength, corporeal and psychological circumstances and psycho-physiological and emotional capacity. Complex operational environment manipulates employees' work performance.

Further, it is significant that staffs are qualified and knows how to perform job related duties with the help of paraphernalia as insufficient utensil usage can affect in mishap or divergences in recital no issue that how a lot and utensil was appropriate. Learning of the individual must also be provided oriented to the good use of defensive tools and individual safety [4].

2.3 Benefits

Benefits can affect the Employee performance in many ways. The benefits are not frequently issued towards assessment and are consequently cheaper to achieve during an employer throughout the market [2]. Therefore, cheaper the benefits ought to amplify employee performance. Further the benefits can proceed as an alternate for salary. To inspect the employer investigation facts and establish the employee's reduced the salary; once numerous benefits had obtained to the individual after a few years [2]. Moreover, the employees sight the benefits and salary as alternatives, disposed to provide salary in trade for extra benefits. While the benefits set as a significant part of employee compensation correspondence and simply acted as organize in different researches, but not as the key issues of analysis.

Furthermore, in the literature the reward shape towards the benefits that individual obtain from their work [11], and the important part of a worker job approach such as business assurance, incentive and employee performance [12]. Thus, in any business rewards in cooperate as a significant role in structuring and supporting the obligation between the workers that guarantee an average job performance and employee loyalty. As, the business entity substitute premise, employee penetrates corporate with the exact set of proficiency towards requiring goals, and anticipate in revisit a respectable working conditions where the employees assist to use their skills and satisfy and attain the desired goals [18]. The rewards enhance the height of effectiveness and efficiency of the individual towards their jobs and in place of outcome there is an improvement of organizational performance [18].

2.4 Empowerment Concept

Empowerment as initially is an important and vital concept by which personnel are capable to effect and have regulator over judgments that influence over their productivity in their job. Moreover, this perception support workers to workout determined regulator or consultant of their effort situation tangibly, communally, ethnically and spiritually over the assets owed to them [5]. Partaking egalitarianism is extra essential idea or belief that qualities enlarging on. Rendering to [25] sharing was careful to be an inherent quality of equality and thus a clarification to suppression. This idea contains the methods and conclusions to inspire development and autonomy.

2.5 Recognition

"Recognition explains as the appreciation to the employees for the level of performance, and success or an influence to achieve goal. It can be intimate or community, fundamental or official. It is continuously in tally to pay." [20]. However the employees also need recognition. Persons like to distribute the celebration of their success with others and have to be recognized in the organization. Whereas the needed is satisfied it works to be an excellent motivator. Further, if employers depend on reward only to recognize influence and success it is most probable that the employee's goals will become altered to protect the pay and nothing more while this will lead to a besmirched culture of the organization. Thus using recognition correctly it will be cost-effective way of increasing success and allow employees to feel intricate in the corporation culture [20].

3 Problem Statement

Today the banks are facing the problem of low productive employees in all the sectors. However, in banking industry the lower level of motivating employees are in number due to the nature of their work. However, this study assist to find out the relationship among the following factors Benefits, Job environment, Empowerment, Recognition, to extract the best known factor among the employees to enhance their motivational level. The study will investigate the impact level for the above mentioned aspects on employee motivation of their job and their after cause on their moral satisfaction and their individual performance in banks.

3.1 Research Question/Objective

The following are the objectives of the research:

- To extract the best known motivational factors that supports the moral intention of individual to perform.
- Which factor is highly motivates employees in the banking industry?
- Do the motivational aspects enhance the individual performance?

4 Hypothesis

4.1 Motivational Factors

$Hypothesis_{a1}(H_{a1})$: There is a significant impact of Benefits over employee Motivation in the Pakistani Banks.
$Hypothesis_{a2}(H_{a2})$: There is a significant impact of Job environment over employee Motivation in the Pakistani Banks.
$Hypothesis_{a3}(H_{a3})$: There is a significant impact of Empowerment over employee Motivation in the Pakistani Banks.
$Hypothesis_{a4}(H_{a4})$: There is a significant impact of Recognition over employee Motivation in the Pakistani Banks.

4.2 Motivation and Individual Performance

$Hypothesis_{a5}(H_{a5})$: There is a significant relationship among motivation and individual performance in the Pakistani Banks.
$Hypothesis_{05}(H_{05})$: There isn't a significant relationship among motivation and individual performance in the Pakistani Banks.

5 Methodology

The research focuses on descriptive research design. Whereas, the precarious connection among the answers about discovering out the connection of Benefits, Job environment, Empowerment, Recognition, that help to enhance the motivational level of individual also to optimize their work performance and as well as its effect on organization. Furthermore explanatory study helps us to explain the relationship among Dependent and independent variables. The descriptive study helps to find out the line of intervention on the factors affecting the individual motivation. However, the job satisfied employee's leads the organization towards the improvement in employee performance and customer satisfaction. This study disquiet with discover how, when, what, why questions concerning about the factors of motivation and clarification of distinct variables may lie in the study which also ornate the modern association among variables that how the delimit factors have their impact over the employee performance.

However, the research depends upon two areas, i.e. the base use which is used for the tactical attitude which trusts over secondary data, comprising seminar proceedings, books, intellectual journals, webinars and articles. Moreover, the research sturdily depends over primary data source by accompanying the survey on employee motivation to collect material about diverse aspects which expressively influence over the employee job performance in banking industry. The banks included in this study are Meezan Bank, Bank Al-Habib, Albaraka, and Faysal Bank. The demographic aspect included age, gender, salary range, job position, service period, and last promotion that received in their organization cover in the questionnaire. Structured Close ended questionnaire is being a part of my study. The questionnaire is the primary tools to gather data for my dissertation. Further human comportment is coherent which rely upon reasons, however all individuals has their own explanation for precise concern they are attaining from convinced area and services even by exploiting certain individual skills, this is the key reason for picking questionnaire as the primary to of data collection.

5.1 Research Model

The motivated employees are the most important assets for the organization. Although it is the considerate perceptional preference of employees being rightly treatment, the functional loom is being centered towards the inference of an employee's motivation on executive success (Fig. 1).

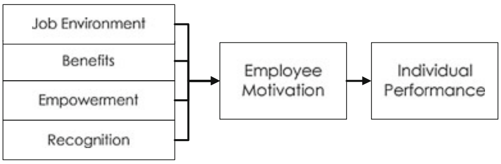


Fig. 1. Conceptual model

Moreover, the employee motivation is in the interest to the executives since it has a reciprocal outcome for every individual and every other thing in organization. Further, in banks employee's motivation is the main factor due to the nature of job and higher employee turnover rate and stress. Further testing will I identify the effect of factors on employee motivation.

6 Results and Discussions

6.1 Employee Preference over the Factors Affects Employee Motivation

The Analysis of the motivational factors that is most preferable by the employees working in the banking sector industry. However, the factor state below indicates that the maximum preferable factor is used to increase the level of individual motivational (Table 1). Further the bar-chart indicate that Benefits share the ratio of 18% to enhance employee motivational level as well as the interpersonal skills of an individual has 13% of its contribution (Fig. 2).

Although the most affective factors that helps to improve the employee inner motivational level is the individual empowerment the employee decision making authority and its contribution among the other factors is almost 22%. However, job environment, recognition and employee growth also have their impact on employee motivational level over their job.

Table 1. Preferable factors that improves employees motivation

Motivational factor	Frequency	Percentage
Benefits	79	26.1589404
Job environment	51	16.88741722
Empowerment	95	31.45695364
Recognition	77	25.49668874
Total	302	100

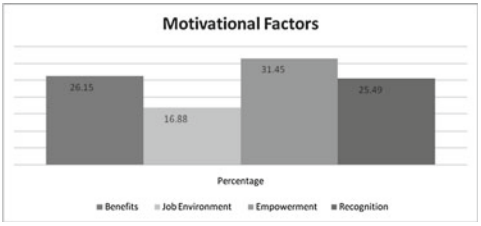


Fig. 2. Preferable factors that improves employees motivation

6.2 Multi Regression Empowerment, Benefits, Job Environment, Recognition and Motivation

The regression equation between empowerment, recognition, growth & development with motivation:

$$\text{Motivation}(y) = f(\text{empowerment, recognition, benefits, job environment}),$$
$$y = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + \text{Error},$$

where b_1x_1 signifies the intercept of empowerment x, b_2x_2 signifies the intercept of recognition x, b_3x_3 signifies the intercept of benefits b_4x_4 signifies the intercept of job environment (Table 2).

Although the table of regression shows the value of R about 0.942 which identifies the strong relationship among independent variable (empowerment, recognition, benefits and job environment) and dependent variable Employee motivation. Whereas, R square shows co-efficient of simple determination which is 88% variance in empowerment, recognition, benefits and job environment explained by motivation. The F-test identifies the rationality of the research model and the model fitness. The significance of F-test is computed where the F calculated is greater than F tabulated which is 0.05. Thus, the null hypothesis H_{01}, H_{02}, H_{03} and H_{04} has been rejected followed by the results and the results also identifies the significant connection between empowerment, recognition, job environment, benefits and employee motivation.

Table 2. Please write your table caption here

Regression Statistics					
Multiple R					0.942
R Square					0.888
Adjusted R Square					0.884
Standard Error					0.1704
Observations					126
ANOVA					
	df	SS	MS	F	Significance
Regression	4	27.92	6.98	240.18	0.000
Residual	121	3.516	0.029		
Total	125	31.437			

6.3 Regression Analysis Between Employee Motivation and Employee Performance

The regression equation between motivation and employee performance:

$$\text{Employee performance}(y) = f(\text{motivation}).$$

The regression outcome identifies the relation among employee motivation and performance which is being further specifies in the below mention table (Table 3).

Table 3. Regression analysis between employee motivation and performance

Regression Statistics					
Multiple R					0.722
R Square					0.522
Adjusted R Square					0.518
Standard Error					0.32622
Observations					126
ANOVA					
	df	SS	MS	F	Significance
Regression	1	14.386	14.386	135.184	0.000
Residual	124	13.196	0.106		
Total	125	27.583			

In Table 3, the outcomes indicates the positive connection among motivation and employee performance is and identify that motivation is also a factor to enhance the employee work performance. However the result shows the regression R 0.722 and adjusted Rsquare is 0.522 which shows the significant relationship. $F_{calculated}$ is greater than $F_{tabulated}$ which is 0.05. Furthermore the result rejects the null hypothesis H_{05}.

6.4 Conclusion and Managerial Implication

The study specifically identifies the importance of motivation in the environment of the banking industries and specifies the factors that improve the employee motivation leveling banks. Whereas, the people spent their life working in the banks, which highlighted the point to improve the motivation level of employees towards their job. This research conducted to keep the motivation level high in banks. Further, the results identify the factors (benefits, recognition, empowerment, and job environment) helps to improve the employee motivation level. Further, each factor shows its level of importance. Moreover the study also explains the satisfactory results from the banks employees about the define factors can enhance employee motivation towards their jobs. The study illustrates that motivated employees satisfied with their job more delicately and serve the organization and customer, especially in the banks where the employees is directly linked with the customer. Whereas the results shows that the factors not only enhance the employee motivation as well as the factors also enhances their moral and social behavior that tends to improve their work performance.

References

1. Ahmad K, Fontaine R (2011) Management from Islamic perspective. Pearson, Malaysia Sdn Bhd (Person custom publishing)
2. Baughman R, Dinardi D, Holtz-Eakin D (2003) Productivity and wage effects of "family-friendly" fringe benefits. Int J Manpow 24(3):247–259
3. Bhatti K (2015) Impact of Islamic piety on workplace deviance. PhD thesis, International Islamic University Kuala Lumpur, Malaysia
4. Bratton J (2012) Human resource management: theory and practice. Palgrave Macmillan, Basingstoke
5. Fetterman DM, Wandersman A (2005) Empowerment evaluation principles in practice. Eval Pract 15:1–15
6. Griffin RW (2012) Management eleventh edition. Cengage Learning, Hampshire
7. Hui MK, Au K, Fock H (2004) Reactions of service employees to organization-customer conflict: a cross-cultural comparison. Int J Res Mark 21(2):107–121
8. Igalens J, Roussel P (1999) A study of the relationships between compensation package, work motivation and job satisfaction. J Organ Behav 20(7):1003–1025
9. Jones G, George J (2011) Essentials of contemporary management 11(2):81–86
10. Jones PH, Chisalita C, Veer GCVD (2005) Cognition, technology, and work: special issue on collaboration in context: cognitive and organizational artefacts. Cogn Technol Work 7(2):70–75
11. Kalleberg AL (1977) Work values and job rewards: a theory of job satisfaction. Am Sociol Rev 42(1):124
12. Kressler H (2004) Motivate and reward: performance appraisal and incentive systems for business success. Palgrave Macmillan, New York
13. Lai C (2009) Motivating employees through incentive programs. Jyvaskyla University of Applied Sciences
14. Leftheriotis I, Giannakos MN (2014) Using social media for work: losing your time or improving your work? Comput Hum Behav 31(31):134–142

15. Levin MA, Hansen JM (2008) Clicking to learn or learning to click: a theoretical and empirical investigation. College Stud J 42:665–674
16. Lin PY (2008) The correlation between management and employee motivation in sasol polypropylene business, South Africa. Polypropylene Business
17. Locke EA, Latham GP (1991) A theory of goal setting and task performance. Acad Manag Rev 15:367–368
18. Mottaz CJ (1988) Determinants of organizational commitment. Hum Relat 41(6):467–482
19. Mullins LJ (2011) Management and organisational behaviour. Oxford University Press, Oxford
20. Robbins SP (2005) Organizational behavior. Prentice Hall, Upper Saddle River
21. Simons T, Enz CA (1995) Motivating hotel employees: beyond the carrot and the stick. Cornell Hospitality Q 36(1):20–27
22. Steers RM, Shapiro DL (2004) The future of work motivation theory. Acad Manag Rev 29(3):379–387
23. Sulaiman M, Ahmad K (2014) The perspective of Muslim employees towards motivation and career success
24. Wang Y (2004) Observations on the organizational commitment of Chinese employees: comparative studies of state-owned enterprises and foreign-invested enterprises. Int J Hum Resour Manag 15(4–5):649–669
25. Wong S, Siu V, Tsang N (1999) The impact of demographic factors on hong kong hotel employees' choice of job-related motivators. Int J Contemp Hospitality Manag 11(5):230–242

How Content Marketing Can Help the Bank Industrial: Experience from Iran

Sima Zomorodian and Yi Lu(✉)

Business School, Sichuan University, Chengdu 610065, People's Republic of China
luyiscu@163.com

Abstract. In today's market, Content marketing is one of the most efficient and successful ways you can do that, but only if you know how to use it correctly. We tried in this article to find, are there any relationship between sensitive content, quality content, new content and content marketing? The study had been designing a questionnaire consists of 37 questions in Likert scale to help us finding our hypotheses clearly, after analyzed the data of questionnaires with SPSS software we fund our three hypotheses and for accept or reject them we use another questionnaires with 40 questions in Likert scale and among 550 randomly selected regular customers of Bank Mellat in Tehran (the capital of Iran) and 400 roperly filled questionnaires were collecting. Then analyzed all the questionnaire with Amos software. The result show all of those items(sensitive content, quality content, new content) can have effect on the customer. So an empirical investigation to determine factors influencing on content marketing in banking industry because in this area there are a lots of brands and customers that can use all of them very easy so they should find different way for attract customers. The result can help the bank industrial to improve their content marketing and increase their customer.

Keywords: Marketing · Content marketing · Export company · Customer

1 Introduction

Content marketing is the process of creating high-quality, valuable content to attract, inform, and engage an audience, while also promoting the brand itself. Content marketing is a kind of Marketing that use words, sounds and photos to improve the knowledge of customer, introduce their brands and new products. Content marketing is about giving away information to build relationships and earn trust, but gating some of your best content is an acceptable and valuable practice. In fact, this is not the only benefit of content marketing, another utility of this kind of marketing is costs less than other types of marketing. Content marketing is a new topic in Iran. Perviously, in bank industrial, they just tried to give better services to customers and hope they introduction their brands. We

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_51

are trying to help the bank industrial for using content marketing to increase their customer. Content marketing is comparable to what media companies do as their core business, except that in place of paid content or sponsorship as a measure of success, brands define success by ultimately selling more products or services [13]. Content marketing is the marketing and Business Process for creating and distributing relevant and valuable content to attract, acquire, and engage a clearly defined and understood target audience - with the objective of driving profitable customer action [10]. "Strong brands are based on a story that communicates who is the company, is to communicate what you really are" [2]. Therefore, the content marketing should be based on the company's values. Also, without considering the quality of the content, which is the most important part of digital marketing, the choice of the frequency of promotion and of the right social media plays a significant role in the success of the content marketing campaign. When building such a strategy, it is important to always have in mind all the social business strategy factors. All content has to work together; all the groups need to work together. There are 6 essential components of content marketing strategy: creation - curation - optimization - social media - amplification - analysis [1]. Content marketing aids in brand recognition, trust, authority, credibility, loyalty and authenticity. Content marketing can help accomplish these tasks for a variety of constituencies, and on several levels: for the organization it represents, for a company's products and services, and for the employee who represent the business or service [8]. In the development of a content marketing, there are numerous opportunities to be more relevant and effective. Planning content that's meaningful to the customers you're trying to engage according of content [9]. Content marketing is anything an individual or an organization creates and/or shares to tell their story. What it isn't: A warmed-over press release served as a blog post. It is conversational, human and doesn't try to constantly sell to you. It also isn't a tactic that you can just turn on and off and hope that will be successful. It has to be a mindset that is embraced and encouraged. You've got to start thinking like a publisher and use that to plan and execute your entire marketing plan which content of any variety should be a part [3]. Making professionally produced creative content available online is proving to be a high-risk business, because of market fragmentation, high development and production costs and the need to fund as yet unprofitable new services from the declining revenue streams of "traditional" analogue and physical distribution [12]. The marketing management philosophy which holds that achieving organizational goals depends on determining the needs and wants of target markets and delivering the desired level of satisfaction more effectively and efficiently than competitors [7]. The forms of content marketing are constantly changing as new tools to create, publish and share that content are launched and others are shut down. Enhancements and new functionality are added to content publishing tools every day, which means the tools you are using to create, publish and share content today might not be the tools you are using tomorrow [5]. The paper has two main research objectives: (1) Study the influential factors on content marketing (2) Study the influential factors on content marketing in Bank Mellat.

The paper is structured as follows: literature review that show our hypothesis come from previous study. Research methodology that show our method for survey the elements. Result that show why we accept hypothesis. Discussion that show the final result and give some suggestion for future research.

2 Literature Review

In today's market, content marketing can solve some problems for company. Actually in Iran this kind of marketing is new and we try in this article to show how this kind of marketing can help to the bank industrial for attract more customers. So we focus on previous studies to find some elements related to content marketing and then use questionnaire to collected data from customers, after that we analyze the data and make the hypothesis. Again we use questionnaire to collected data to accept or reject the hypothesis. Content marketing is defined as a marketing process of creating and properly distributing the content in order to attract, make communication with, and understand other people so that they can be motivated to do beneficial activities [4]. More specific solutions are outlined in content marketing strategies 'where I focus on quality of the content, its creation, distribution and evaluation' [1]. In addition to monitoring mentions and share, engaging with people who responded to the content can be a very powerful way to spread your reach and to connect with potential prospects or industry stakeholders [10]. The explanation is that "When a brand uses specific words or stories that resonate with a consumer, they can dig deeper into who they are as a consumer. By utilizing content marketing, brands can cater campaigns and stories around buying patterns and personalities" [2]. Base on those researches we make this hypothesis: Quality of content has effect on export companies customers. To be relevant to your audience and create a powerful brand you must win their trust and admiration. With the creation of valuable content you build interest that transforms into lasting relationships [11]. Due to the characteristics of these emerging technologies, the digital content market is growing rapidly and traditional content providers face service transformation decisions. While a majority of the previous technology adoption studies have focused on the viewpoints of users and customers, cost reduction, or electronic channel related technologies, in this research we analyze the emerging technology adoption decisions of competing firms for providing new content services from a strategic perspective. Utilizing game theoretical models, we examine the effects of market environments (technology cost, channel cannibalization, brand power, brand extension, information asymmetry and market uncertainty) on firms' adoption decisions [6]. Base on those researches we make this hypothesis: Sensitive of content has effect on export companies customers. However, by implicitly restricting it's focus to a buyer's perspective, the resource-based and relational views also leave the question of which resources and competencies add greater value-for-customer largely unanswered [14]. According to some authors, the contribution of different benefits to a relationship's value will vary along the relationship life-cycle. Specifically, the supplier's know-how of the supply market, the adaptation and improvement of extant products and the development of new products become increasingly relevant in explaining the value of a relationship from the customer's perspective as the life-cycle advances [14]. Base on those researches we make this hypothesis: New of content has effect on export companies customers.

3 Research Methodology

For finding the elements related to content marketing in the bank industrial, we find some elements from previous studies use questionnaire in the different branch of Mellat bank and the customer answer them after that we analyze the data. The questionnaire consists of 40 questions divided into two sections: the first, consisting of 4 multiple choice questions, allow us to obtain general information about the sample (gender, age, years of service and Education level). The second, measured by a Cronbach's Alpha analyze:

- Has quality of content effect on bank industrial customers?
- Has sensitive content effect on bank industrial customers?
- Has new content effect on bank industrial customers?

The area of this research is Tehran. The capital of Iran. Statistical Society of this research is the customers of Mellat bank in 2016. Scale of five options to describe the items listed in the questionnaire of "very large impact" to "very little impact" is used. A total of 370 questionnaires were collected at the end of the questionnaire, as shown in Table 1.

Table 1. The elements

Element	Questionnaires	The percentage	
New content sharing	0.623	Less than 0.001	Accept
Brand	0.742	Less than 0.001	Accept
Content creation	0.701	Less than 0.001	Accept

4 Result

In the first analysis, factor analysis was exploratory and in the second one confirmatory. There are plans to test certain hidden factors beyond the variables. In this case, orders are expected variables. In the researchers to test hypotheses related to a particular factor structure of the action. In this analysis, the researchers tried to obtain empirical data based on a model that assumes a relatively small number of parameters, describing explain or justify. Reliability Statistics: as shown in Table 2.

KMO and Bartlett's Test: as shown in Table 3.

The SPSS software for data Factor and AMOS software for factor analysis was used, the result is as shown in Fig. 1.

Table 2. Reliability Statistics

Cronbach's Alpha	N of items
.909	28

Table 3. KMO and Bartlett's Test

Kaiser-Meyer-Olkin measure of sampling adequacy	0.895
Approx Chi-Square	2076.635
Bartlett's test of sphericity df	190
Sig	0.0000

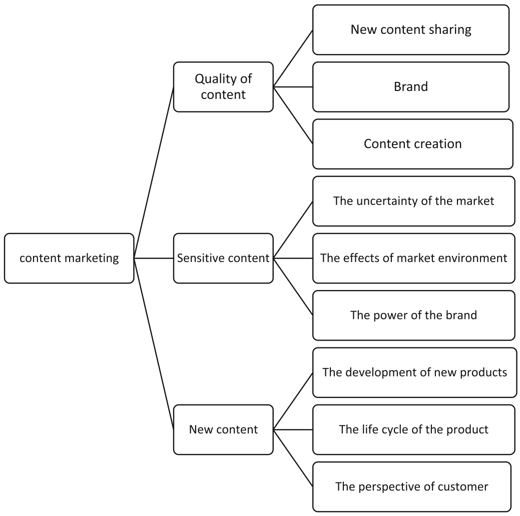


Fig. 1. Factor analysis result

(1) Quality of Content
Exploratory factor analysis in the first order for quality of content, and finally, the three most important variable component factors identified in accordance with the following table can be significant.as shown in Table 4.

(2) Sensitive Content
Exploratory factor analysis in the first order for sensitive content, and finally, the three most important variable component factors identified in accordance with the following table can be significant, as shown in Table 5.

(3) New Content
Exploratory factor analysis in the first order for new content, and finally, the three most important variable component factors identified in accordance with the following table can be significant, as shown in Table 6.

Table 4. Quality of content

The index	Standard coefficient	Significant level	Acceptance or rejection situation
New content sharing	0.623	Less than 0.001	Accept
Brand	0.742	Less than 0.001	Accept
Content creation	0.701	Less than 0.001	Accept

Table 5. Sensitive content

The index	Standard coefficient	Significant level	Acceptance or rejection situation
The uncertainty of the market	0.586	Less than 0.001	Accept
The effects of market environment	0.492	Less than 0.001	Accept
The power of the brand	0.606	Less than 0.001	Accept

5 Discussion

In content marketing the customers are not presented with a direct suggestion, instead they take another issue and information that is useful for them. At the end they trust the content provider. Content marketing also can use with another kinds of marketing.

In this paper, we examine customer opinion to analyze some elements that can be effective on content marketing. In Iran this kind of marketing is new so we try to show the bank industrial that they can use it to improve their business.

Our results show that there are elements related to content marketing. However, they are sensitive content, quality of content and new content. In conclusion, In today's marketing, content marketing due to lower fee than other marketing methods can be very effective. In the banking industry in Iran due to the existence of different banks and brand diversity is very high, customers in the choice of the Bank are having doubts.

The most simple and low-cost way to attract customers in order to help create content provides the services of the Bank and the Bank's competitive advantages compared to other banks to inform customers. In this context must be carefully that creating content is purely informational aspect and do not have directly tried to attract customers, because it may have a reverse and does not help to attract customers.

Table 6. New content

The index	Standard coefficient	Significant level	Acceptance or rejection situation
The development of new products	0.545	Less than 0.001	Accept
The life cycle of the product	0.714	Less than 0.001	Accept
The perspective of customer	0.810	Less than 0.001	Accept

After content attract trust of a group of customers, this group begin to do mouth to mouth advertising, this action does not have charge for Bank. According to the result of the study, New content can improve the knowledge of customer and can help them to make a good decision. Must be careful in creating content do not use imitation method because the duplicate content to help not only our business, but does not cause a lack of quality in the customer's mentality.

To create the content need a team to be active and knowledgeable in this field, that continuously target market observation and based on customer needs to start creating content.

Content marketing is the present and especially the future of marketing and inherently to marketing in general. Therefore, a content marketing strategy cannot be successful without having a quality of content. Higher quality of content created as compared to other competitors is the number of people who can trust our business will increase. In today's world most people unwilling to spend time for unnecessary information. So the creation of quality content is one of the critical components in creating the content. Also when we improve the quality of content we can make a competitive advantage and this is improve our chance to have more customer. In the uncertainty market we should be careful about which kind of content is necessary for customer.

In creating the content should be think about specific time like New year. Create the appropriate content at any time, can be very useful in the success of the business. So teams that create content need to be aware of the point and make a suitable content for customers.

Content marketing is still in developmental stage. There is a need to consider content for bank's industrial. There is an opportunity for banks to use content marketing more and more for finding new customer.

For future research there are some suggestions. We can do this research on another industrial. Also evaluation of the factors influencing the content marketing customer in a bank-centered and compare it to other banks. Factors influencing the relation between detection of content-driven marketing and customer satisfaction.

Acknowledgements. This research is supported by the Program of Social Science Foundation of Sichuan (Grant No. SC16C010). The authors are also grateful to Sichuan University for providing research funding (No. 2016SCU11036; No. skqy201640).

References

1. Augustini M (2014) Social media and content marketing as a part of an effective online marketing strategy. PhD thesis, Doctoral dissertation, Diploma thesis, Masaryk University: Faculty of informatics, Brno
2. Baltes LP (2015) Content marketing-the fundamental tool of digital marketing. Bull Transilvania Univ Brasov Econ Sci Ser V 8(2):111
3. Elisa R, Gordini N (2014) Content marketing metrics: theoretical aspects and empirical evidence. Eur Sci J 10(34):92–104

4. Forouzandeh S, Soltanpanah H, Sheikhahmadi A (2014) Content marketing through data mining on facebook social network. Webology 11(1):1
5. Gunelius S (2011) Content marketing for dummies. Wiley, New York
6. Jin BH, Li YM (2012) Analysis of emerging technology adoption for the digital content market. Inf Technol Manage 13(3):149–165
7. Kotler P (2009) Marketing management: a South Asian perspective. Pearson Education India
8. Lieb R (2011) Content marketing: think like a publisher-How to use content to market online and in social media. Que Publishing, Upper Saddle River
9. Odden L (2012) Optimize: how to attract and engage more customers by integrating SEO, social media, and content marketing. Wiley, Hoboken
10. Patrutiu Baltes L (2015) Content marketing-the fundamental tool of digital marketing. Bulletin of the Transilvania University of Brasov Series V: Economic Sciences 8(2):111–118
11. Peltekoğlu DFB, Peltekoğlu YI (2016) The digitalization of communications: towards content marketing. Temmuz, pp 776–780
12. Rosati E (2013) Originality in EU copyright: full harmonization through case law. Edward Elgar Publishing, Cheltenham
13. Ström C, Pöysti S et al (2015) Developing an event marketing strategy using digital media: case: data group jyväskylä
14. Zerbini F, Golfetto F, Gibbert M (2007) Marketing of competence: exploring the resource-based content of value-for-customers through a case study analysis. Ind Mark Manage 36(6):784–798

The Establishment and Application of AHP-BP Neural Network Model for Entrepreneurial Project Selection

Ke Wu[1](✉) and Xiaofeng Li[2]

[1] Department of Enterprise Management, School of Business, Sichuan University, Chengdu 610065, People's Republic of China
wuke@scu.edu.cn

[2] Department of Management Science, School of Business, Sichuan University, Chengdu 610065, People's Republic of China

Abstract. It is even the most difficult and critical step in starting a business. Unfortunately, prior entrepreneurial project selection (EPS) methods are barely satisfactory. In this paper, we constructed AHP-BP neural network mode based on the traditional BP neural network. It effectively reduces the input dimensions of traditional BP neural network, significantly raises its learning speed, and improves the prediction accuracy. In the end, empirical study shows that this model is able to fully absorb the tacit knowledge of expert assessment, reduce subjective randomness, which provides important references to entrepreneurs in starting a business, thus it is worthy being popularized and applied.

Keywords: Entrepreneurial project · Analytic hierarchy process · BP neural network · Risk decision

1 Introduction

Entrepreneurial activities have shown increasingly close relationship with job creation in both developed and developing countries, and have become a strong driving force to world's economic development. Meanwhile, entrepreneurship has occupied a prominent position on the agendas of policymakers and researchers [14]. In complex environments, Triple Helix has been operationalized in different ways, spaces, and contexts where those agents are transforming their roles in the development and strengthening of national innovation and entrepreneurial ecosystems [3]. Unfortunately, at early start, failure is always going along with those entrepreneurial projects. According to 2015 Annual Report of GEM (Global Entrepreneurship Watch), the huge success made by 20%–30% startups in China is at the cost of 70–80% failures. Even in developed countries, 35% startups failed in the first year, almost half of them disappeared in five years [15]. Numerous reasons exist but mainly because entrepreneur started with an inappropriate project. For instance, Nabil provided a starting point for a stronger

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_52

theoretical grounding of research that goes beyond the traditional interpretation of entrepreneurial failure and opens new avenues of research opportunities to explore and compare the different configurations that have emerged and identify the possible dynamism and trajectories among these configurations [7]. However, how to use scientific methods to select appropriate projects is particularly important for entrepreneurs [11].

Currently, scholars are dedicated to studying the theory and method of EPS, and made some achievements [4,6,8,12,19,21]. These studies could be divided into two categories: first, qualitative analyses based on theories, including SWOT analysis, brainstorming and vision method. These studies has certain feasibility but underper forms in practice. Secondly, quantitative analyses using mathematical models, including income constraint method, utility function method, risk ranking method and fuzzy evaluation method. These methods are barely satisfactory in practice because of subjective factors. Therefore, how to choose a right entrepreneurial project has become a difficult problem in management.

This paper is organized as follows. In Sect. 2, the evaluation index system of EPS is built. In Sect. 3, the traditional BP artificial neural network is introduced. In Sect. 4, we combined AHP and BP neural network to construct the AHP-BP neural network model of EPS. In Sect. 5, the empirical research based on 21 entrepreneurial projects shows that the model is scientific and practical.

2 Evaluation Index System

Entrepreneurial Project Selection (EPS) means choosing the right entrepreneurial projects at the beginning of business. There are a lot of factors influencing EPS, subjectively and objectively. Thus, we need to follow certain principles and methods when establishing scientific index system of EPS. The principles are as follows:

(1) Goal oriented principles. The index system should be designed according to the needs of entrepreneurs, and keep in line with policy guidance, and it should be able to evaluate the market prospects of the entrepreneurial project.
(2) Scientific principle. Evaluation indicators must be able to reflect and evaluate the entrepreneurial projects correctly.
(3) Systematic principle. The indicators should be able to reflect the horizontal and vertical relationship between different levels.
(4) Operational principles. The selected indexes should be easy to practice. The index system should be simplified, be able to provide qualitative and quantitative evaluation to projects.

According to the principles above, the factors that influence the normal operation of the enterprise are classified into several parts. By sending questionnaires, interviewing experts and referring to literature at home and abroad [2,17,20,22], the frame is subdivided, supplemented and deleted in detail. Finally, the index system is built, as is shown in Table 1.

Table 1. The evaluation index system of entrepreneurial project

Target layer A	Criterion layer (weight)	Index layer	Weight of Index	Comprehensive weight
The evaluation index system of entrepreneurial project	Environmental competitiveness (0.066)	Industry environment	0.3946	0.026
		Industrial policy	0.2119	0.014
		Legal environment	0.2043	0.0135
		Regional advantage	0.1892	0.0125
	Technical competitiveness (0.463)	Advanced technology	0.3946	0.1827
		Technical complexity	0.2449	0.1134
		Technical reliability	0.1374	0.0636
		Technical substitutability	0.2082	0.0964
		Technical accumulation degree	0.015	0.0069
	Market competitiveness (0.095)	Market demand	0.3187	0.0303
		Market share	0.0579	0.0055
		Competitor capability	0.1749	0.0166
		Market entry barriers	0.2843	0.027
		Cost-effective of products	0.1642	0.0156
	Financial competitiveness (0.067)	Capital requirement	0.2362	0.0158
		Financial strength	0.1378	0.0092
		Fund-raising ability	0.3422	0.0229
		Internal rate of return	0.2044	0.0137
		Financial management ability	0.0795	0.0053
	Internal competitiveness (0.309)	Practical experience	0.1074	0.0332
		Leadership ability	0.3311	0.1023
		Entrepreneurial skills	0.254	0.0785
		Entrepreneurial knowledge	0.1223	0.0378
		Social resources	0.1853	0.0573

3 BP Neural Network Architecture and Algorithm

3.1 Overview of AHP Methods

The analytic hierarchy process (AHP), is formally proposed by operational research experts Thomas L. Satty (T.L. Satty) in the middle of 1970s [9]. AHP decomposes a complex problem into various components, and group these factors into an orderly hierarchical structure. The relative importance of the factors is determined by pair wise comparison, and the overall order of the relative importance is determined by the comprehensive human judgment. The appearance of AHP brings great convenience to the decision makers in solving the decision problem which is difficult to be quantitatively described, it can be used in almost any scientific field. The detailed calculation steps of AHP method can be found in literature [1]. But Analytic Hierarchy Process, which analyzes complex decisions by organizing the problems into a multilayer hierarchic structure, is a simple yet popular decision technique used extensively in every decision field and is inadequate to handle the uncertain decision making problems [13].

3.2 Algorithm of Traditional BP Neural Network

BP (Back Propagation) neural network is a kind of error back propagation training algorithm for the multilayer feed forward network, it is one of the most widely used neural network model [10], proposed by a group of scientists led by Rumelhart and McCelland in 1986. BP neural network reflects many basic features of human brain. It is a highly complex nonlinear dynamic learning system, especially suitable for processing imprecise and fuzzy information.

Traditional BP neural network is a kind of non feedback feed forward network [18]. It is composed of input layer, hidden layer and output layer. A typical topology of three layer forward BP network is shown in Fig. 1. The first layer consists of n input nodes. Each of the n input nodes is connect to each of the r nodes in the hidden layer. The r output nodes of the hidden layer are all connected to each of the m node in the output layer.

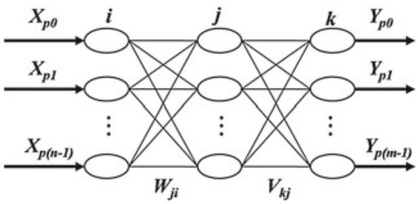


Fig. 1. Typical three layers BP network structure

The connection weights of input nodes i and hidden nodes j is w_{ij}, the connection weights of hidden node j and the output node k is v_{kj}, the threshold of the hidden layer node is, the threshold of the output layer node is θ_j. The model is $\theta_k^{'}$ provided with N learning samples $(X_P, Y_P)(p = 1, 2, \cdots, N)$. $X_P = (x_{p0}, x_{p1}, \cdots, x_{p(n-1)})^T$ is the input vectors for learning sample P, $Y_P = (y_{p0}, y_{p1}, \cdots, y_{p(n-1)})^T$ is the ideal output vector for learning sample P.

In input layer, let the input equals output, that is:

$$O_{pi} = x_{pi}(i = 0, 1, \cdots, n-1). \tag{1}$$

The operating characteristic of the middle hidden layer node is:

$$\mathrm{net}_{pj} = \sum_{i-0}^{n-1} w_{ji} O_{pi} - \theta_j, \tag{2}$$

$$O_{pj} = f(\mathrm{net}_{Rj})(j = 0, 1, \cdots, r-1). \tag{3}$$

The operating characteristic of the output layer node is:

$$net_{pk} = \sum_{j-0}^{r-1} v_{kj} O_{pj} - \theta_k^{'}, \tag{4}$$

$$O_{pk} = f(\mathrm{net}_{pk})(k = 0, 1, \cdots, m-1). \tag{5}$$

O_{pi}, O_{pj}, O_{pk} are the output of input layer node, hidden layer node and output layer node separately. net_{pj}, net_{pk} are the input of hidden layer and output layer separately. Sigmoid function is used as transfer function $f(\cdot)$: $f(x) = 1/(1+e^{-x})$.

For convenience, in the Eqs. (2), (4), let:

$$w_{jn} = -\theta_j, O_{pn} = 1, v_{kr} = -\theta_k^{'}, O_{pr} = 1.$$

Equations (2) and (4) become:

$$\text{net}_{pj} = \sum_{i-0}^{n} w_{ji} O_{pi} \quad O_{pj} = f(\text{net}_{pj}) = \frac{1}{1+e^{-\text{net}_{pj}}}, \tag{6}$$

$$\text{net}_{pk} = \sum_{j-0}^{r} v_{kj} O_{pj} \quad O_{pk} = f(\text{net}_{pk}) = \frac{1}{1+e^{-\text{net}_{pk}}}. \tag{7}$$

The learning process of the input and output samples of the traditional BP network is the adjustment of error between the actual output and sample output of network weights, it is actually a nonlinear optimization problem uses the most common gradient descent method in mathematical programming:

$$w(n+1) = w(n) - b \cdot \frac{\partial E}{\partial w} \quad v(n+1) = v(n) - b \cdot \frac{\partial E}{\partial v}.$$

In the above equity, b refers to learning rate. In the traditional BP algorithm, b is constant. The value of b directly affects the convergence and convergence rate of the algorithm.

3.3 Improvement of Traditional BP Neural Network Algorithm

Traditional BP neural network has many disadvantages. The steps of the traditional BP neural network which was improved by HE Chang-zheng etc. [5] are as follows:

Step 1. Based on actual problem, set the number of input and output layer nodes, and initialize the parameters, including the learning accuracy ε, the prescribed number of iterative steps M_0, he upper limit of hidden layer nodes R, learning parameters b, momentum coefficient a, the initial number of R is 1. The admissible region of the initial weight H is divided into N equal parts $\beta_1 \beta_2 \cdots \beta_N$.

Step 2. Input the learning samples and make the values of sample parameters [0, 1].

Step 3. Give random numbers between −1 and 1 to the initial weighting matrix $\beta_i(\beta_i \subset H)(i = 1, 2, \cdots, N)$.

Step 4. Train the network with the traditional BP method.

Step 5. Judge whether or not the number of iterative steps is exceeding M_0. If yes, continue. If no, turn to Step 7.

Step 6. If $r+1$ exceed the upper limit of R, if yes, end, no, turn to Step 3.
Step 7. Determine whether the accuracy of learning meets the requirements, if yes, continue, if no, turn to Step 4.
Step 8. Call the deletion and merge algorithm of hidden nodes, if there is a deletion or merger, turn to Step 4, otherwise, algorithm completes.

4 BP-AHP Model Construction

Table 1 indicates that 24 factors affect the choice of entrepreneurial projects. By using all of them as the input variables, the complexity will obviously increase and we may lose the calculation precision. Fortunately, AHP could address this issue. Firstly, acquire the important evaluation indicators by AHP, and use them as input variables of the BP network. Secondly, train the improved BP algorithm. When fitting with other conventional methods, the accuracy of AHP is often not as good as the BP algorithm. Therefore, we combine these two methods organically, so as to enhance the problem solving ability of artificial neural network in the complex nonlinear problem.

4.1 Determine the Input Variables of BP Network Using AHP

Based on the analysis of the factors influencing EPS, a multi-level hierarchical structure can be established (see Table 1). Then, construct the judgment matrix and determine the weight of the index. Invite experts to conduct in-depth analysis and reach a consensus on the importance of each index, use the $9-1$ scale methods to establish the various layers of λ_{max} of each judgment matrix and its corresponding characteristic vector are obtained respectively. Finally, the feature vectors are normalized to get the comprehensive weight of each evaluation index, see Table 1, $CR = 0.016 < 0.10$, consistency test passed.

The results show that important indicators affect the choice of entrepreneurial projects are: Industry Environment (x_1), Advanced Technology (x_2), Technical Complexity (x_3), Leadership Ability (x_4), Technology Substitutability (x_5), Entrepreneurial Skills (x_6), Technical Reliability (x_7), Social Resources (x_8), Entrepreneurial Knowledge (x_9), Practical Experience (x_{10}), Market Demand (x_{11}), Market Entry Barriers (x_{12}), Fund-Raising Ability (x_{13}), Competitor Capability (x_{14}), Capital Requirement (x_{15}). The above 15 indicators have 90% or more impact on EPS; they are input variables of the BP network. Other indicators are less influential, can be deleted.

Denote $V_j (j = 1, 2, \cdots, n)$, $V = \{V_1, V_2, \cdots, V_n\}$ represents the comment set of entrepreneurial projects. V = excellent, fine, ordinary, bad $=$ $A1, A2, A3, A4 = 1, 2, 3, 4$.

In order to evaluate the entrepreneurial projects objectively, we set up a team to select evaluation indicators of entrepreneurial projects. The team members include 3 entrepreneurs, 3 venture capitalists, 3 business management experts. The experts involved in the evaluation are required to grade the indexes. Then, denote $v_k = \frac{1}{9}\sum_{i-1}^{9} vu_i$, v_k stands for the evaluation value of index, x_k,

$i (i = 1, 2, \cdots, 9)$ stands for the 9 experts. vu_1 stands for the grade given by experts, $u_1 \in [0, 1]$ stands for "poor", $u_2 \in (1, 2]$ stands for "bad", by analogy, $u_5 \in (4, 5]$ stands for "excellent".

4.2 Construction of BP-AHP Model

According to the structure of BP neural network, the BP-AHP model includes three layers:

(1) Input Layer. First of all, assign the evaluation of the index selected by the AHP method; and they are used as the input layer of the neural network variables. From what has been stated above, 15 input layer variables belong in the neural network.
(2) Hidden Layer. According to the improved learning algorithm of BP neural network proposed by TIAN Jing-wen etc. [16], set the number of hidden layer nodes as 1, then, the network will learn by itself until we get the appropriate number of nodes.
(3) Output Layer. The comments of entrepreneurial project evaluation are defined as two categories: "choose it" and "give it up", these can be represented by the output vectors $(1, 0)$, $(0, 1)$. So, the number of output layer nodes is 2.

AHP and BP neural network methods are combined to create the BP-AHP model of EPS, and the basic steps of the algorithm are as follows:

Step 1. Simplify the evaluation index of EPS, and the important index is selected as the input layer of neural network.
Step 2. Set the number of neural network output layer nodes, and initialize the network parameters (including the given learning accuracy ε, provisions iteration number M0, the limit of hidden nodes r, learning parameter η. The initial hidden node number is set to 1.)
Step 3. Input the learning samples and make the values of sample parameters $[0, 1]$.
Step 4. Give random numbers between -1 and 1 to the initial weighting matrix.
Step 5. Train the network with the modified BP method, and determine the weight matrix between each layer.
Step 6. Judge whether or not the number of iterative steps is exceeding the prescribed. If yes, end; If no, go back to step 5 and continue learning.
Step 7. Collect values of the indexes and process these data to make them $[0, 1]$.
Step 8. Input the processed data to the trained BP neural network and get the output.
Step 9. According to the output results, combined with the entrepreneurial project index reviews set, choose the right entrepreneurial projects.

5 Empirical Research

We choose 21 entrepreneurial cases from Sichuan University in the past 3 years, of which 10 projects have been successful, 11 entrepreneurial projects ended in failure. Use the BP-AHP model to evaluate entrepreneurial projects selected. The cases are: computer training (Q_1), online digital products (Q_2), development and production of management software (Q_3), computer DIY and maintenance (Q_4), website construction and service (Q_5), computer cleaning services (Q_6), water saving car washing (Q_7), reward supervision social software (Q_8), online customized travel (Q_9), advertisement designing (Q_{10}), anime 3D printing (Q_{11}), a beverage shop (Q_{12}), fruit fast delivery (Q_{13}), creative souvenirs design (Q_{14}), fast dry cleaning service (Q_{15}), household service (Q_{16}), cozy cafe (Q_{17}), creative home decoration (Q_{18}), children's art training (Q_{19}), architectural design (Q_{20}), online second hand market (Q_{21}). Q_1, Q_2, Q_3, Q_5, Q_6, Q_{10}, Q_{12}, Q_{16}, Q_{17}, Q_{19}, Q_{21} failed in the first half year; others operate smoothly. We used the first 16 entrepreneurial projects as training samples of the BP - AHP model, the last 5 items as prediction evaluation samples.

5.1 BP-AHP Neural Network Training

In the process of quantitative evaluation of the selection of venture capital projects, 9 experts were invited to set up the quantitative evaluation group of the index system. According to the evaluation index of the evaluation team, the evaluation team can be used to evaluate the evaluation index of the entrepreneurial project 1 to the venture capital project (x_k), and the evaluation value of the evaluation index x_k is obtained

$$v_k = \frac{1}{9}\sum_{i-1}^{9} vu_i.$$

The output of successful projects are $(1, 0)$, the output of failed projects are $(0, 1)$. The BP neural network architecture is built by 15-1-2 (The number of input layer nodes is 15, the number of hidden layer nodes is 1, the number of output layer nodes is 2). We initialize the network (the upper limit $\varepsilon = 0.0002$, learning rate $\eta = 0.5$, Inertia Parameter $a = 0.1$), Make $x_k(k = 1, 2, \cdots, 15) \in [0, 1]$ by dividing them with 10, and input the processed data to the BP network model. Then the network is trained by the modified BP learning algorithm and the network architecture becomes 15-9-2. At the same time, we get the optimized network weight matrix.

5.2 Entrepreneurial Project Selection

Now, it is time to evaluate Q_{17}, Q_{18}, Q_{19}, Q_{20}, and Q_{21} with the trained neural network. Give values to the indexes x_k of the 5 entrepreneurial projects, Make $x_k \in [0, 1]$ by dividing them with 10. Input the processed data to the trained

BP neural network and get the outputs as shown in the second row of Table 2. According to the principle of maximum membership degree, we can determine the evaluation results of each project; see the third row of Table 2.

Table 2. BP neural network output

Entrepreneurial projects	BP neural network output	Evaluation result	The actual operation of the business project
Q17	(0.0129, 0.9965)	Give it up	Failed
Q18	(0.9847, 0.0355)	Choose it	Succeed
Q19	(0.0287, 1.0002)	Give it up	Failed
Q20	(0.9798, 0.0032)	Choose it	Succeed
Q21	(0.0195, 0.9921)	Give it up	Failed

It can be seen from Table 2, the projects which evaluation is "Give it up" failed in actual operation; the "Choose it" ones succeed. The predictions of BP neural network completely accord with the practical ones, which indicate that, this risk early warning model feasible and effective.

6 Conclusions

EPS is a comprehensive prediction problem affected by many factors, so it is important to choose appropriate evaluation method. According to the characteristics of entrepreneurial projects, this paper analyzes the factors that affect EPS, and constructs the evaluation model of EPS based on AHP and BP neural network subjectively and objectively. This model can not only extract the main attributes of entrepreneurial projects, reduce the input variables of the neural network, but also reduce the complexity of the neural network and the training time, improve the generalization ability, reasoning ability and classification ability. Therefore, the model is feasible and effective, and it has important reference value in guiding the entrepreneur to choose the right project.

Acknowledgments. This research was supported by the National Social Science Foundation of China (No. 16BGL024).

References

1. Brennan MJ, Schwartz ES (1985) Evaluating natural resource investments. J Bus 58(2):135–157
2. Deng Z, Yang P, Yang X (2015) The way of college students' entrepreneurial project selection. China Manag Informationization 18(21):235–236 (in Chinese)
3. Guerrero M, Urbano D (2016) The impact of triple helix agents on entrepreneurial innovations' performance: an inside look at enterprises located in an emerging economy. Technol Forecast Soc Change

4. Hagan MT, Demuth HB, Beale M (1996) Neural Network Design. PWS Publishing Company, Thomson Learning, Boston
5. He C, Li X, Yu H (2002) The new improvement of bp neural network and its application. Math Pract Theory 32(4):555–561
6. Jain R, Ali SW (2013) A review of facilitators, barriers and gateways to entrepreneurship: direction for future research. South Asian J Manage 20:122–163
7. Khelil N (2015) The many faces of entrepreneurial failure: insights from an empirical taxonomy. J Bus Ventur 31(1):72–94
8. Kumaraswamy MM, Rahman MM (2002) Risk management trends in the construction industry: moving towards joint risk management. Eng Constr Architectural Manage 9(2):131–151
9. Lockett G, Stratford M et al (1986) Modelling a research portfolio using ahp-a group decision process. R&D Manage 16(2):151–160
10. Lu X (2010) The main strategies for college students to choose the entrepreneurial projects. Innov Entrepreneurship Educ 6:21–23 (in Chinese)
11. Nielsen SL, Lassen AH (2011) Images of entrepreneurship: towards a new categorization of entrepreneurship. Int Entrepreneurship Manage J 8(1):35–53
12. Pena I (2002) Intellectual captital and business venture success. J Intellect Capital 3:180–198
13. Ren P, Xu Z, Liao H (2016) Intuitionistic multiplicative analytic hierarchy process in group decision making. Comput Ind Eng 101:513–524
14. Schelfhout W, Bruggeman K, Maeyer SD (2016) Evaluation of entrepreneurial competence through scaled behavioural indicators: validation of an instrument. Stud Educ Eval 51:29–41
15. Schwab K (2015) The Global Competitiveness Report, 2014–2015. World Economic Forum, Geneva
16. Tian J, Gao M (2006) Artificial Neural Network Algorithm Research and Application. Beijing Institute of Technology Press (in Chinese)
17. Urban B, Nikolov K (2013) Sustainable corporate entrepreneurship initiatives: a risk and reward analysis. Technol Econ Dev Econ 19:383–408
18. Wen L, Li L, Liao S (2013) Analysis on influencing factors of college students' entrepreneurial project selection. Project Manage Technol 11(6):89–92 (in Chinese)
19. Xie Q, Yuan X (2006) A review of the evaluation system of venture investment projects. Sci Technol Manage Res 8:182–187 (in Chinese)
20. Yan Y (2011) Application of fuzzy comprehensive evaluation method in the selection of college students' entrepreneurial projects. Project Manage Technol 9(12):51–55 (in Chinese)
21. Zhang N, Yan P (1998) Neural Networks and Fuzzy Control, 1st edn. Tsinghua University Press, Beijing (in Chinese)
22. Zhu Y (2014) Selection of college students' entrepreneurial projects from the perspective of the long tail theory. Innov Entrepreneurship Manage 5(6):52–54 (in Chinese)

Improving the User Experience and Virality of Tourism-Related Facebook Pages

Ayako Sawada[1(✉)] and Taketoshi Yoshida[2]

[1] Hokuriku Gakuin Junior College, Ishikawa, Japan
sawada@hokurikugakuin.ac.jp
[2] Japan Advanced Institute of Science and Technology, Ishikawa, Japan

Abstract. The percentage of people using social network services (SNSs) exceeds 60% in Japan; SNSs are being used by people of various age groups. Recently, the number of companies that use Facebook Pages to advertise their activities and the number of organizations that use them to disseminate information have been increasing. In tourism, disseminating local information on Facebook Pages and increasing the attractiveness of the pages are expected to lead to the promotion of regional tourism industries. In this study, the current state of the use of tourism-related Facebook Pages was examined. The engagement rate and the number of fans were used as an indicator of the attractiveness of Facebook Pages. An increasing the attractiveness of Facebook Pages was examined through emotion analysis of user comments.

Keywords: Facebook page · Regional development · User comments · Emotion analysis

1 Introduction

Along with the spread of smartphones, the number of people using social network services (SNSs) has been increasing. The number of monthly active users of Facebook worldwide (those who use Facebook one or more times per month) reached 1.79 billion as of 30 September 2016 [1]. In addition, it is reported that the number of monthly active users of Instagram is 0.6 billion (as of 15 December 2016) [4], that of Twitter is 0.313 billion (as of 30 June 2016) [12], and that of Line is 0.2184 billion (as of 31 March 2016) [6]. According to the 2015 White Paper on Information and Communications in Japan published by the Ministry of Internal Affairs and Communications [5], the SNSs used in Japan over the last one year include Line (37.5%), Facebook (35.3%), and Twitter (31.0%). The number of people using SNSs generally tends to decrease with increasing age: approximately 50% of people aged 20 or younger, less than 40% of people in their 30's and 40's, and more than 20% of people aged 60 or older use Facebook. It is reported that Facebook is being used by people of various age groups. With this background, a variety of research projects on Facebook have been

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_53

carried out. Wilson et al. classified 412 sociology papers on Facebook into five categories: descriptive analysis of users, motivations for using Facebook, identity presentation, the role of Facebook in social interaction, and private information disclosure [13].

Recently, both the number of companies that use Facebook Pages to advertise their activities and the number of organizations that use them to disseminate information have increased. Ohara et al. clarified the characteristics of photos that increase the responses of users of Facebook Pages of fast-food companies by principal component analysis [8]. In tourism, disseminating local information on Facebook Pages and increasing the attractiveness of pages are expected to lead to the promotion of regional tourism industries. By multiple regression analysis, Sabate et al. investigated the factors that increase the attractiveness of Facebook Fan Pages (currently, Facebook Pages) of five travel agencies in Spain using the numbers of "Likes" and comments regarding each post as indices of attractiveness [9]. The results indicate that the number of "Likes" is affected by the presence of video images and photos in the post and that the number of comments is affected by the presence of photos and the time period of posting. However, the methods of using Facebook are rapidly advancing and most recent posts almost always include photos or video images.

Research on Facebook and its usefulness for those involved in the promotion of regional tourism industries is limited, although how-to books on posting on Facebook Pages [3] and marketing strategies using Facebook [10] have been published. Consequently, Sawada et al. [11] analyzed the content of tourism-related posts on Facebook Pages with high attractiveness and clarified the following. The responses of users of Facebook Pages offering information related to Christmas, flowers, and foods were high, whereas those for pages offering information on cultural events tended to be low. The responses of users of Facebook Pages offering real-time information, such as information on the start of the blooming of cherry blossoms, tended to be high. In addition, it was found that the responses of users were affected by the type of photos posted.

To increase the attractiveness of Facebook Pages, the analysis of not only the posts of organizations but also the feelings of users that cause them to post comments need to be clarified. The purpose of this study is to analyze the current state of the use of Facebook Pages in tourism in Japan and the sentiments behind the comments on Facebook Pages with high attractiveness, and to clarify the feelings of the users who posted comments and the factors behind the increase in the attractiveness of these pages.

In Sect. 2, the current state of the use of Facebook Pages offering regional tourism information is analyzed in terms of the number of fans, management, and the responses of users to find the characteristics of Facebook Pages with high attractiveness. In Sect. 3, user comments on these Facebook Pages are analyzed. In Sect. 4, the factors behind the increase in the attractiveness of pages are discussed on the basis of the results of analyzing user comments. Section 5 is the conclusion.

2 Facebook Pages in Tourism

Facebook navi (http://f-navigation.jp/) is the only navigation site in Japan recognized by Facebook, Inc. In this study, 842 Facebook Pages registered in the Facebook Page ranking in the travel category of Facebook navi were examined. The data were collected on 10 January 2016. In this section, Facebook Pages offering regional tourism information were selected and the relationship between the number of fans and organization managing the site was clarified. Then, Facebook Pages with high attractiveness were determined on the basis of the responses of users and the number of fans.

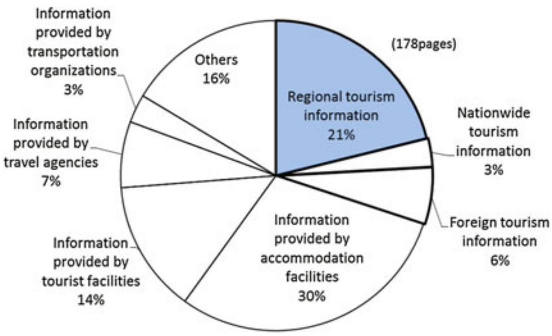


Fig. 1. Content of Facebook Pages in the travel category

2.1 Contents of Facebook Pages

The authors browsed the 842 target pages in the travel category and classified them depending on their contents. Figure 1 shows the results. Among the 842 pages, 30% of the pages are managed by accommodation facilities, 14% by tourist facilities such as aquariums and museums, 7% by travel agencies, and 3% by transportation organizations. Some of these pages offer regional tourism information such as festivals, events, and the start of the autumn foliage season. However, these pages are basically used as a public relations tool of these facilities and organizations. Other pages disseminate regional tourism information rather than advertising the activities of organizations. The management of these pages varies from companies and individuals to volunteer groups. Twenty-one percent of the pages (178 pages) focus on information about particular regions; 3% focus on information about specific topics such as one-day hot spring trips, famous places for flowers, and Buddhist statues; and 6% focus on information about foreign countries. In this study, we targeted the 178 pages focusing on information about particular regions.

2.2 Pages Offering Regional Tourism Information

(1) Number of Fans
The users clicking the "Like" button of a Facebook Page are registered as fans of the page. The postings on the page appear in the News Feed of the user, and the user can browse the various types of information on a Facebook Page on which they are registered as fans. The number of users who have clicked the "Like" button is "the number of fans". Figure 2 shows the number of fans of the 178 Facebook Pages offering regional tourism information. The number of fans of the top-ranking page exceeds 70,000. There are 19 pages with more than 10,000 fans. There are 143 pages with less than 5,000 fans, and 72 of the 143 pages with less than 1,000 fans. For most of the pages, the number of fans is limited and the number of users browsing the pages offering regional tourism information is low, suggesting that the effect of these pages on the promotion of regional tourism industries is limited.

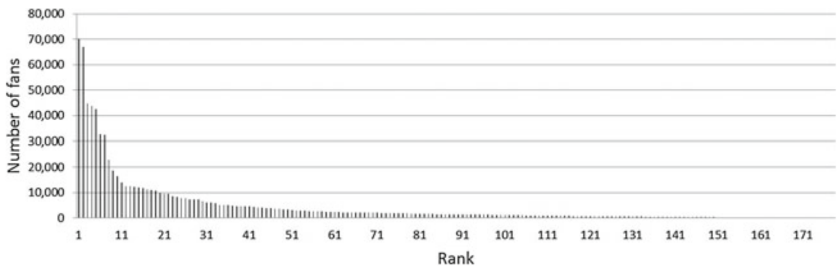


Fig. 2. Number of fans of the 178 Facebook Pages offering regional tourism information

(2) Management
Figure 3 shows the classification of the management of the 178 pages offering regional tourism information. Among the pages, 36, 24, and 9% of the Facebook Pages are managed by individuals, private companies, and local governments, respectively. Others include regional tourist associations in the form of general incorporated associations, incorporated nonprofit organizations (NPOs), and volunteer groups.

Figure 4 shows the classification of the management of Facebook Pages with more than 10,000 fans among the 178 pages. Considering all the Facebook Pages offering regional tourism information, more than 30% of the pages are managed by individuals, as shown in Fig. 3; however, only 11% (2 pages) of Facebook Pages with more than 10,000 fans are managed by individuals, as shown in Fig. 4.

An organization that manages one of the Facebook Pages with more than 40,000 fans was interviewed on 30 October 2015. The management consisted of a total of nine members including one young staff member and eight people who provide regional information. They have adopted the management rule that the frequency of posting is one or two posts per day. The interviewee said that finding

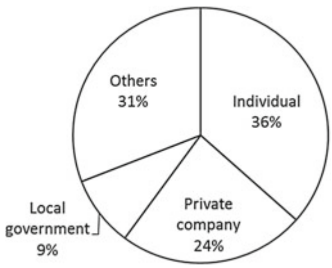


Fig. 3. Management of Facebook Pages offering regional tourism information

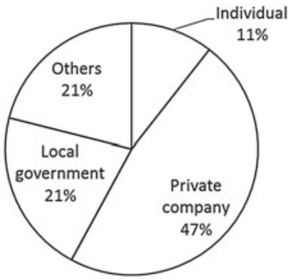


Fig. 4. Management of Facebook Pages with more than 10,000 fans

new material every day is troublesome work even though there are nine members. When a Facebook Page is managed by individuals, finding raw material and collecting information require much time and expense, leading to difficulty in posting at a high frequency.

(3) Facebook Pages with High Attractiveness
Table 1 shows a summary of 19 Facebook Pages with more than 10,000 fans. Facebook Pages for various areas from Hokkaido to Okinawa attract many fans, although 2 pages on Yokohama are also included in the list. For these 19 Facebook Pages, posts, the presence or absence of photo and video images, the presence or absence of links to other homepages, the number of "Likes" for each post, the number of "Shares" of each post, and the number of user comments were obtained between January 1 and December 31, 2015, by Graph API.

Engagement rate is an index for evaluating the attractiveness of pages on SNSs. The conventional engagement rate of Facebook Pages is given by

$$\text{Conventional engagement rate} = \begin{aligned}&(\text{Number of “Likes”} + \text{Number of}\\ &\text{comments} + \text{Number of “Shares” of}\\ &\text{the post})/\text{Number of fans.}\end{aligned} \tag{1}$$

However, conventional engagement rate defined in this way decreases as the number of fans increases [2]. Therefore, Facebook, Inc. defined a new engagement rate as

$$\text{New engagement rate} = \begin{aligned}&(\text{Number of “Likes”} + \text{Number of comments}\\ &+ \text{Number of “Shares”} + \text{Number of people who}\\ &\text{have clicked on the post})/\text{Number of reaches.}\end{aligned} \tag{2}$$

This was released in January 2015 or later but the exact date of introduction was not publicized. The new engagement rate cannot be calculated by third parties, other than the management, because the reach count and click count of the page are used in the calculation. Therefore, in this study, the conventional engagement rate calculated using Eq. (1) was used to determine the engagement rate per fan in Table 1. According to the Facebook Engagement Survey 2014 [2], the mean engagement rate of Facebook Pages with $\leq 10,000$ and $< 50,000$ fans is 1.996% and that of Facebook Pages with $\leq 50,000$ and $< 100,000$ fans is 1.499%. The mean engagement rate of Facebook Pages in the travel/leisure category is 1.56% (mean number of fans: 30,642).

The number of Facebook Pages with engagement rates of $< 1\%$, ≥ 1 and $< 2\%$, ≥ 2 and $< 3\%$, ≥ 3 and $< 4\%$, and ≥ 8 and $< 9\%$ per fan is 5, 2, 4, 7, and 1, respectively. The Facebook Page on Shirakawa-go accepts posts from users, and the number of posts is extremely high. The engagement rate per post is small, as is the engagement rate per fan.

In this study, both the engagement rate per fan and the number of fans are used as indices of attractiveness. Three Facebook Pages with an engagement rate of $\geq 3\%$ and a large number of fans, namely, the Facebook Pages of "Kyushu Tourism Information", "Akita Vision", and "Yokohama Tourism Information", were selected from Table 1 as Facebook Pages with high attractiveness to be analyzed.

3 Analysis of User Comments

In this section, the user comments on the pages with high attractiveness were classified into ten types to clarify the feelings that cause users to post comments.

3.1 Analysis Method

For the Facebook Pages of "Kyushu Tourism Information", "Akita Vision", and "Yokohama Tourism Information", user comments collected over a period of one year (2015) were analyzed to avoid seasonal bias.

First, engagement rates of all the posts for the three pages in 2015 were calculated. We read all the user comments for the posts with the top 100 engagement rates (hereafter, user comments for the top 100 posts) and the user comments for the posts with the lowest 100 engagement rates (hereafter, user comments for the lowest 100 posts) and classified the user comments in accordance with the feelings that caused users to post comments. For this classification, ten feelings in the *Dictionaryon Expression of Feeling* [7]; namely, "joy (pleasure)", "anger (unpleasantness)", "sadness (crying)", "fear (apprehension)", "shame (humiliation)", "liking (longing, nostalgia)", "disgust (loathing, regret)", "sensation (inspiring, excitement)", "ease (calm)", and "surprise (amazement)". Table 2 shows examples of user comments classified into these ten feelings. Examples 1–9 are classified as one of the ten feelings. When multiple feelings are included in one comment, as in example 10, the user comment is classified as multiple feelings. Comments including the intension of "share" in the sentence are classified as "share". The comments that do not belong to any of the ten feelings are classified as "others". Stamps without any comments are classified as "stamp".

Table 1. Facebook Pages with more than 10,000 fans

Rank	Facebook page name	Management	Number of fans	Engagements number			Engagement rate		Number of posts between Jan. 1-Dec. 31 2015
				"Likes"	"Shares"	Comments	Per post	Per fan	
1	Kyushu Tourism Information	General incorporated association	70,105	856,559	38,515	5,493	2,427	3.5%	371
2	Shirakawa-go	Private company	67,049	1,125,034	44,538	6,147	450	0.7%	2,612
3	Akita Vision	Local government	44,780	621,495	26,001	4,637	1,513	3.4%	431
4	The Heart of Osaka-Visit Osaka Japan	Local government	43,673	27,622	553	151	44	0.1%	651
5	Yokohama Tourism Information	Public interest incorporated foundation	42,655	651,163	23,136	3,370	1,433	3.4%	473
6	Yokohama China Town	Cooperative association	32,748	149,274	4,468	1,054	499	1.5%	310
7	Gooood Place!! in Shiga	Private company	32,566	236,962	7,080	1,297	1,141	3.5%	215
8	Otaru Fan	Private company	22,830	31,289	27	9	870	3.8%	36
9	Exchange of Kyoto Information -Kyoto Now-	Private company	18,592	99,287	5,473	436	408	2.2%	258
10	Hyogo Tourism Guide	Local government	16,470	18,832	529	115	81	0.5%	239
11	Okinawa Diving	Private company	13,870	1,520	18	13	111	0.8%	14
12	Okayama Great Spot Net	Public interest incorporated association	12,547	76,720	3,081	721	323	2.6%	249
13	Fukuoka No Machi	Individual	12,472	4,241	51	28	103	0.8%	42
14	I LOVE TOKUSHIMA	Individual	12,267	62,946	3,432	797	1,050	8.6%	64
15	Hokkaido Fan Magazine	Private company	12,053	179,615	10,050	1,129	305	2.5%	625
16	Let's Visit Nara	Private company	11,857	40,560	1,590	157	188	1.6%	225
17	Kamakura Block	Private company	11,275	58,557	1,301	671	348	3.1%	174
18	Web magazine [Shikoku Tairiku]	Private company	11,026	58,428	2,468	362	286	2.6%	214
19	Nagasaki Tourism Promotion Section	Local government	10,657	58,284	2,647	384	383	3.6%	160

3.2 Analysis Results

Figures 5 and 6 show the classifications of feelings underlying the user comments for the top 100 posts and those underlying the user comments for the lowest 100 posts, respectively. The total number of user comments for the top 100 posts classified as one of the feelings, "share", "stamp", or "others" is 6,924. The percentage of "liking (longing, nostalgia)" was the highest (25.9%), followed by "sensation (inspiring, excitement)" (17.9%). The total number of user comments for the lowest 100 posts classified as one of the feelings, "share", "stamp", or

Table 2. Examples of classifications of feeling of user comments

	User comment	Classification
Example 1	I will go there next week (^^). I can't wait	Joy
Example 2	It's a waste of electricity	Anger
Example 3	I can't visit to take photos because of work this year (T_T)	Sadness
Example 4	I'm worried about fainting because of the high temperature	Fear
Example 5	I love Yokohama!	Liking
Example 6	Oh, my god. I can't take holidays during silver week because of work	Disgust
Example 7	It's a beautiful night scene!!!	Sensation
Example 8	I feel relaxed. I'll enjoy the night scene when I go back to my hometown	Ease
Example 9	5000, wonderful!	Surprise
Example 10	It's really beautiful!! I'd like to visit there and ride it!	Sensation, Liking
Example 11	I will share the page	"Share"
Example 12	Can I get it in Tokyo?	"Others"

"others" is 1,200. The percentage of "liking (longing, nostalgia)" was the highest (29.8%), followed by "share" (11.3%) and "sensation" (8.2%). There are no user comments for either the top or the lowest 100 posts classified as "shame".

The numbers of the user comments for the top and lowest 100 posts in each feeling classification were compared. As shown in Table 3, we developed a 2×2 contingency table to summarize the number of user comments classified as one of the feelings and the total number of user comments classified as other feelings, "share", "stamp", and "others", to carry out Fisher's exact test. A contingency table was developed for each feeling in the test. Table 4 shows the summary of the data and results. A significant difference was observed between the user comments for the top and lowest 100 posts for six feelings, "share", "stamp", and "others". The underlines in Table 4 indicate which of the top or lowest 100 posts has a higher number of user comments than the expected value for the classification with significant difference.

3.3 Tendency of User Comments Classified as Each Feeling

The number of user comments for the top and lowest 100 posts classified as "joy" is small. Although a significant difference was observed between the two, the content of comments is an expression of gratitude, such as "thank you for the information on xxx", "I'm looking forward to attending xxx" for the announcement of an event, and "I'm looking forward to eating xxx" for seasonal food.

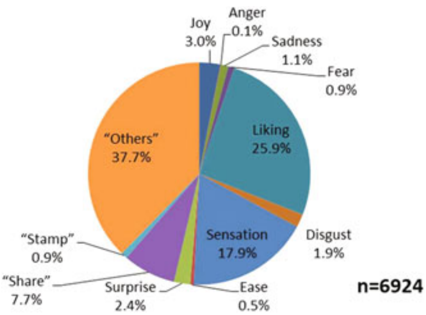


Fig. 5. Classification of feelings of user comments for the top 100 posts.

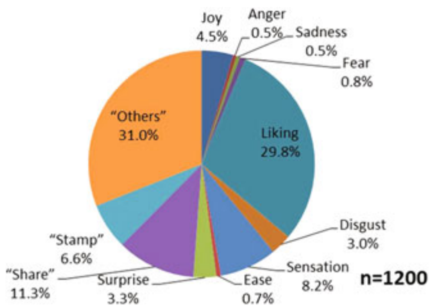


Fig. 6. Classification of feelings of user comments for the lowest 100 posts.

Table 3. Contingency table summarizing the number of particular feelings for the user comments for the top and lowest 100 posts.

	Top	Lowest
Feeling X	a	b
Feelings other than feeling X	c	d

a: The number of user comments for the top 100 posts classified as feeling X
b: The number of user comments for the lowest 100 posts classified as feeling X
c: The number of user comments for the top 100 posts classified as feeling other than X
d: The number of user comments for the lowest 100 posts classified as feeling other than X

The number of user comments for the top and lowest 100 posts classified as "anger" is small. Although a significant difference was observed between the two, there are few user comments expressing anger and unpleasantness.

The number of user comments for the lowest 100 post classified as "sadness" is small. For the top 100 posts, most of the user comments classified as "sadness" express sadness that they cannot visit a place because of work despite a desire to do so or that photos of their hometown cause a longing to return.

The number of user comments for the lowest 100 posts classified as "fear" is small. Examples of the user comments for the top 100 posts include concern regarding going to a place introduced in the post and the fear of unconventional designs and ideas in the post.

There are no user comments classified as "shame" for either the top or the lowest 100 posts.

The number of user comments classified as "liking" is the largest among the user comments for both the top and lowest 100 posts. Although a significant difference was observed between the two, the content of comments is an expression of liking, such as "I like xxx hot spring" and "I love xxx!", the expression

Table 4. Number of comments for the top and lowest 100 posts classified into each feeling category and the results of Fisher's exact test

Feeling	Top	Lowest	p value of Fisher's exact test
Joy	211	54	0.013*
Anger	5	6	0.002*
Sadness	78	6	0.045*
Fear	59	10	1.000
Liking	1793	357	0.006*
Disgust	133	36	0.021*
Sensation	1237	98	0.000*
Ease	35	8	0.515
Surprise	165	39	0.088
"Share"	533	135	0.000*
"Stamp"	62	79	0.000*
"Others"	2613	372	0.000*

Fisher's exact test: * $p < 0.05$.
The underlines indicate which of the top or lowest 100 posts has a greater number of user comments than the expected value. Also, a significant difference was observed between the user comments for the top and lowest 100 posts.

of longing, such as "I would like to visit xxx once in the future", and the feeling of nostalgia because of overlapping old memories with the post. People often express liking when they read posts related to their past experience and knowledge, such as a previously eaten food, a place they would like to visit in the future, and memories of childhood.

The user comments for the top and lowest 100 post classified as "disgust" mostly express feelings of regret, such as "I would like to go but I cannot". Only a few user comments express undesirable and unfavorable feelings.

The number of user comments for the top 100 posts classified as "sensation" tended to be large, and a significant difference was observed between users comments for the top and lowest 100 posts. Among the user comments for the top 100 posts, the number of user comments classified as "sensation" is the second most numerous after those classified as "liking". Users are moved by photos of beautiful scenes and excited by photos of delicious-looking food for both categories of posts.

The number of user comments for the top and lowest 100 posts classified as "ease" is small. Users feel relief or healed upon viewing the photos posted.

Regarding "surprise", for user comments on the top 100 posts, users are surprised at the scale and impact of photos or amazed by the beauty of photos. User comments for the lowest 100 posts express surprise because they found unexpected content in the post, such as "I never imagined that such a place existed there".

Some users making comments classified as "share" feel empathy with the post and express their intent to share the post. However, users simply click the share button (a simple comment expressing the intension as is often the case). A significant difference was observed between the user comments for the top and lowest 100 posts. The number of user comments classified as "share" is greater for the lowest 100 posts. We request that users related to the management and the fans of the Facebook Page share posts regardless of their feelings toward the contents of posts.

The number of user comments for the top and lowest 100 posts classified as "stamp" is small. A significant difference was observed between the two, and the number of user comments for the lowest 100 posts classified as "stamp" tended to be larger.

The number of user comments classified as "others" is largest for both the top and lowest 100 posts. Examples of user comments include questions, such as example 12 in Table 2, reports of past experience, such as "I have visited there" and "I ate it yesterday", and expressions of intent and plans, such as "I will visit there next" and "I will go to eat there again this year". We suggest that the comments reporting past experiences are triggered by the feeling of liking the post. The comments expressing intent and plans are possibly triggered by the feeling of sensation (users are moved or excited about the post) and the feeling of joy. However, we cannot detect the true feeling of users, and we classified user comments that do not include the direct expression of feelings as "others" in this study.

4 Discussion

In this section, directions for increasing the attractiveness of Facebook Pages for those involved in the promotion of regional tourism industries are discussed by referring to actual examples of user comments for which a large number of user comments expressed the feelings of users.

4.1 User Comments for the Top 100 Posts Classified as Liking

Example 1 is a post related to seasonal fish. Users report that the fish is delicious and describe the characteristic taste of the fish during breeding season. Example 2 is a post reporting the arrival of the season for tourist boats with kotatsu, a small table having an electric heater underneath and covered by a quilt. Comments 2-1 and 2-2 indicate that users have enjoyed this type of boat before, recommend it to others, and express their intent to experience it again. The user of comment 2-3 had experienced the boat without kotatsu and expresses a willingness to experience the boat with kotatsu this season.

Among the user comments for the top 100 posts, those classified as "liking" are the most numerous. To increase the attractiveness of Facebook Pages for tourist spots and food, it is indispensable to offer posts of famous tourist spots and food that many users like, want to visit, and experience. For well-known information (tourism resources), timely posts considering the best season to enjoy the resources will contribute to the improvement of the engagement rate through links between the information and the lifestyle and memories of users.

Example 1. Seasonal fish

- Comment 1-1 ("liking") I ate that fish. It was delicious!
- Comment 1-2 ("liking") I've eaten this fish! It had many eggs and was delicious.

Example 2. Cruise on a boat with kotatsu

- Comment 2-1 ("liking") The boat with kotatsu is only in winter and was quaint. It was really nice (^_^)/!
- Comment 2-2 ("liking") We ate steamed eel after enjoying a boat with kotatsu!! I'd like to visit again.
- Comment 2-3 ("liking") I visited there about 5 years ago because it's my husband's hometown(^o^)/. It was December, but the boat didn't have a kotatsu. I'd like to try it when I visit again(o^^o).

4.2 User Comments for the Top 100 Posts Classified as Sensation

Example 3 is a post related to the illumination of the buildings and gardens of temples and shrines of a certain region. Users posted comments 3-1 and 3-2 because they were moved by the beautiful photos, although it is not clear whether the users actually had visited the place or not. The user posting comment 3-3 was moved by the beautiful photos and posted a comment although the user had not visited the place.

The user comments classified as "liking" were mostly posted in response to objects that users already knew, whereas the user comments classified as "sensation" were posted by users who were moved or excited by the photos regardless of their previous knowledge. This finding suggests that the content and quality of the attached photos are important and that posting of photos that may induce sensation and excitement in users is effective when posting information (tourism resources) on sites that people may not be aware of.

Example 3. Illumination of buildings and gardens of temples and shrines

- Comment 3-1 ("sensation") It's really beautiful!
- Comment 3-2 ("sensation") The illumination of a temple in Fukuoka prefecture is beautiful.
- Comment 3-3 ("sensation") It's wonderful. I've been there once before, but I didn't know that it was such a beautiful and wonderful place.

4.3 User Comments for the Lowest 100 Posts Classified as Liking

Example 4 is a post related to local sweets. From comments 4-1 and 4-2, we can see that users feel nostalgia for their favorite sweets. In contrast, comments 4-3 and 4-4 indicate that the sweets are not famous; only a small number of people know about the sweets. Example 5 is a post related to a local restaurant. From the post and comments 5-1 and 5-2, it is clear that the users know the restaurant introduced in the post and agree on its attractiveness. From comments 5-3 and 5-4, it is clear that the users would like to visit the restaurant and eat the dishes described in the post.

Among the user comments for the lowest 100 posts, the number of comments classified as "liking" was the largest (357). The users who have eaten the sweets or visited the restaurant before, as shown in Examples 4 and 5, are thought to feel empathy for the posts and make comments because of a feeling of nostalgia and the memory of the food and service, although users who did not know the contents of the posts showed no interest in and did not respond to the posts. In addition, users related to the region are interested in the posts because information new to them is provided in the posts. This finding suggests that the posts about places and foods familiar to those from the region or to the local people, even though not famous, can attract a certain amount of responses from users, leading to an increase in the attractiveness of Facebook Pages.

Example 4. Local sweets

- Comment 4-1 ("liking") I love the sweets! This reminds me of the sweets made by my mother.
- Comment 4-2 ("liking") It brings back old memories(*^ ^*). I want to eat it again!
- Comment 4-3 ("others") I grew up in Akita, but don't know these sweets.
- Comment 4-4 ("others") I don't know them, either.

Example 5. Local restaurant

- Comment 5-1 ("liking") I've frequently visited this restaurant! The hamburger, rice omelet, and demiglace sauce there are delicious!
- Comment 5-2 ("liking") I loved the dishes (≥∇≤)
- Comment 5-3 ("liking") They look delicious. I want to eat them.
- Comment 5-4 ("liking") I'd like to visit this restaurant!

5 Conclusion

In this study, we carried out an analysis of the feelings expressed in user comments on Facebook Pages with many fans offering regional tourism information, and clarified the feelings of the users who posted the comments and the factors behind the increase in the attractiveness of these Facebook Pages.

Compared with websites, Facebook Pages can be easily started and have no startup costs. Therefore, the number of pages managed by individuals is large. When focusing on Facebook Pages with many fans, the percentage of Facebook Pages managed by individuals is small; the Facebook Pages managed by organizations such as regional tourist associations and companies offering regional tourism information are the mainstream. However, such pages are not always actively accessed by many visitors. The information explained in this paper may be useful for increasing the attractiveness of Facebook Pages.

The engagement rate of the Facebook Page "I Love Tokushima" (Table 1) is extremely high (8.6%). The number of posts per year is 64, which means an annual posting frequency of 1.2. According to Locowise Ltd. (the UK), the engagement rate of Facebook Pages with 1-4 posts per week is the highest. The survey by Locowise Ltd. targeted all Facebook Pages, not limited to those of the tourism industry. In future studies, we will investigate the relationship between the frequency of posting and the engagement rate of Facebook Pages on tourism to determine the optimal frequency of posting.

Acknowledgments. This study was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Number 15K00476).

References

1. Facebook (2017) Statistics of facebook. http://newsroom.fb.com/company-info/
2. Facenavi (2017) Facebook engagement survey 2014. http://facebook.boo.jp/facebook-engagement-survey-2014
3. Inc Comnico (2014) Getting Fans! Know-How of Posting on Facebook. Shoeisha
4. Instagram (2017) Statistics of instagram. https://www.instagram.com/press/
5. Ministry of Internal Affairs and Communications (2017) 2015 white paper on information and communications in japan. http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h27/pdf/
6. Line (2017) January-march 2016 business report. https://linecorp.com/ja/pr/news/ja/2016/1347
7. Nakamura A (1993) Dictionary on Expression of Feelings. Books Tokyodo, Tokyo
8. Ohara S, Kajiyama T, Ouchi N (2015) Extraction of characteristics of photos to increase the engagement rate of facebook pages of companies analysis of photos of products of fast-food companies. IEICE Trans A J98–A(1):41–50
9. Sabate F, Berbegal-Mirabent J, Cañabate A et al (2014) Factors influencing popularity of branded content in facebook fan pages. Eur Manag J 32(6):1001–1011
10. Saito S (2014) Facebook Marketing [Business Technique] Technique of Obtaining Valuable "Like". Shoeisha
11. Sawada A, Yoshida T, Murakami K (2016) Factors behind increasing attractiveness of tourism-related facebook pages. In: Proceedings for the 13th Workshop of the Society for Tourism Informatics, pp 17–20
12. Twitter (2017) Twitter usage/company facts. https://about.twitter.com/ja/company
13. Wilson RE, Gosling SD, Graham LT (2012) A review of facebook research in the social sciences. Perspect Psychol Sci 7(3):203–220

Meta-analysis of the Factors Influencing the Employees' Creative Performance

Yang Xu[1], Ying Li[1](✉), Hari Nugroho[2], John Thomas Delaney[3], and Ping Luo[1]

[1] Business School, Sichuan University, Chengdu 610064, People's Republic of China
liyinggs@scu.edu.cn

[2] Department of Sociology, Gedung F. Faculty of Social and Political Sciences, Universitas Indonesia, Depok 16424, Indonesia

[3] Kogod School of Business, American University, 4400 Massachusetts Ave, Washington, DC 20016-8044, USA

Abstract. Innovation has become a popular topic in society. The creative performance fluctuates with the employees' mood caused by the working environment. Many factors have been suggested to be affecting the employees' creative performance. The paper answered the questions as "What are the factors that affect the employees' creative performance through employee behavior? What is the proportion (weight) of each of these determinants?" Furthermore, this study resolves this problem via the meta-analysis, which refers to do systematic quantitative analysis by former studies to summarize the research results of creative performance over the last ten years, and to ascertain the factors influencing the creative performance as well as to determine the correlation between them. Suggestion having been proposed including the further investigation of the interaction between the factors impacting the employees' creative performance as well as explore the corresponding effects. And it is necessary to note that changing one factor may eliminate the effect of several factors, which may cause negative change in the staff creative performance and there is joint efforts needed in terms of enhancing the employees' creative performance.

Keywords: Factors · Creative performance · Meta-analysis

1 Introduction

Creative performance refers to the products, processes, methods, and ideas created by the employees, which are novel, practical, and valuable to an organization. Although creativity varying from person to person cannot be generalized, it is not fixed. In other words, creativity is constantly fluctuating. Likewise, the creative performance fluctuates with the employees' mood caused by the working environment. What are the factors affecting the employees' creative performance? Do they have positive or negative impacts on the staff? Which of these factors has more profound influence and which has less? These issues are

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_54

considered by both business managers and numerous scholars. By surveying past scholarship on creative performance, the paper finds that there are similarities and differences between the researches. Specifically, some examine the impact of internal factors on the employees' creative performance; some explore the external factors; some investigate the employees' creative performance affected by both internal and external determinants, while others study several other influencing factors. All these researches, however, are not comprehensive. What are the factors that affect the employees' creative performance through employee behavior? What is the proportion (weight) of each of these determinants? The paper resolves this problem via the meta-analysis, hoping to summarize the research results of creative performance over the last ten years, and to ascertain the factors influencing the creative performance as well as to determine the correlation between them.

1.1 Related Concepts

Creative performance refers to the products, processes, methods and ideas created by the employees, which are novel, practical, and valuable to an organization. Work passion, a core characteristic of self-identification, refers an employee's strong love for a certain occupation, which he or she considers vital and into which this individual tends to invest time and effort.

Autonomous orientation refers to an individual's desire to act, based on his or her interests and self-values. More precisely, this individual considers incidents as opportunities and challenges, and he or she will take active actions to seize these opportunities and to shoulder the responsibilities, showing strong internal and extrinsic motivations.

Control orientation refers to an individual's desire to act, based on other individuals' attitudes rather than his or her own ideas, because he or she is vulnerable to the external environment.

Mastery approach goal orientation refers to the development process in which an individual tends to concentrate on the abilities, learning, and tasks. Performance approach goal orientation refers to an individual's emphasis of the performance-related matters.

Performance avoidance goal orientation refers to an individual's avoidance of the performance-related matters.

Cognitive re-evaluation refers to an individual's control of the emotional response through changing the impact of events on personal awareness.

Inhibition of expression refers to an individual's control of the emotional response by suppressing the emotional expression.

Feedback consistency refers to the consistency between an employee's expectation of the boss's feedback on his or her work performance and the actual feedback this individual receives.

1.2 Literature Review

(1) Overview of Foreign Researches

As early as the 20th century, the academic circle started researching on the creativity and creative performance. Analyzing the impact of leadership style and support on the employees' creative performance through empirical studies, Huang et al. [4], argued that charismatic and supportive leaders will facilitate the creative performance. Based on the investigation of the relationship between the employers and the employees, their research results provide references for the managers. The above researchers, however, do not consider the impacts of the employees' personal factors on the creative performance. Ulger [7] examined the impact of employees' participation in the creative processes and the pleasure and challenge of creative tasks on their creative performance through empirical researches. They claim that the more staff participation and the more intriguing and challenging the task, the better the creative performance will be. Moreover, the time of the staff participation process, during which multiple factors may contribute to the change of staff creativity, will determine the level of the final creative performance. Nevertheless, these scholars leave the second point for future research. Chong and Ma [1] explored the influence of the personal factors of employees, such as motivation, mood and demand, on the creative performance via empirical researches, he suggest that the motivations and demands of employees generally have positive impacts on their creative performance. Furthermore, in terms of mood, the negative emotions have an adverse effect on their creative performance and vice versa. Their researches assist readers to understand what affects the employees' creative performance from the internal perspective, but they fail to consider the external factor.

(2) Overview of Domestic Researches

Likewise, being interested in this topic, Chinese researchers conduct a host of studies. Using the empirical research, Zhang [8], Zhang et al. [9], Zhang et al. [10], Tang et al. [6], examine the impacts of organizational environment (for instance, working time and pressure, organizational situation, and leadership) on the employees' creative performance. They conclude that the staff creativity is positively correlated with the working time and pressure as well as leadership styles; the influence of organizational situations varies from person to person. Song [5], Zhang et al. [9], Dong et al. [2], analyzing the impacts of the employees' emotions on their creative performance, concur that the negative emotions have adverse impacts on creative performance, and vice versa. Due to other factors in these studies, their research data differ moderately. Similar to foreign scholars, the above researchers only investigate one aspect that affects the staff creativity. Gong et al. [3], exploring the influence of employees' motivations on their creative performance via the empirical research and scenario analysis, ascertain that the creative performance is inhibited by the workers' extrinsic motivations and control motivations but promoted by the employees' other motives. Their final results differ because of various classifications of the motivations.

Overall, Chinese and foreign scholars have investigated numerous factors affecting the creative performance, and their conclusions, which differ slightly in each essay, are roughly the same. In terms of research methods, these papers, however, principally adopt empirical research and scenario simulation. As for the investigations of the influencing factors, they largely highlight one of them. Based on the literature above, this paper utilizes a different approach– the meta-analysis, seeking to examine the factors that impact the employees' creative performance holistically.

1.3 Introduction to the Meta-analysis

(1) Definition and Characteristics of the Meta-analysis

The Meta-analysis is widely used in disciplines, such as medicine, psychology, economics, and management. In these fields, numerous scholars, for the most part, study the same problem, but the final results may be different due to the process control, error, and sample selections. Therefore, they fail to reach a clear conclusion. It is of necessity and significance to analyze the results achieved by other scholars through the Meta-analysis.

Meta-analysis was first proposed by American scholar Glass when he gave a speech at the American Educational Research Association in 1976. According to Glass, meta-analysis refers to a statistical analysis that combines the results of multiple individual studied to resolve uncertainty when reports disagree.

Nonetheless, there is no unified definition of the Meta-analysis. This paper adopts the definition from the book titled Conducting Meta-Analysis through Stata: Meta-analysis is a statistical approach that examines the results of the individual studies of the same topic, analyzing the similarities and the sources of difference between these research outcomes. Meta-analysis and systematic evaluation are different concepts, even though there are similarities between them. Systematic evaluation refers to a rigorous evaluation and review of an issue via a certain search strategy, in the process of which the Meta-analysis may or may not be used.

There are some salient features (advantages) of the Meta-analysis, compared with the individual study: (i) to improve the statistical efficacy; (ii) to evaluate the results quantitatively and comprehensively, resolving uncertainty when reports disagree; (iii) to evaluate the hypothesis, discovering the deficiencies of the previous researches; (iv) to analyze numerous literature data simultaneously, without being affected by the number of references; (v) Using the individual research results as the original data, Meta-analysis is a higher level of logical analysis.

(2) Advantages and Disadvantages of the Meta-analysis

Each research method has its advantages and limitations, and the paper presents a brief introduction to the merits and demerits of the Meta-analysis.

The Meta-analysis is used extensively because of four advantages. First, the Meta-analysis is more accurate than the significance test in terms of the estimation and reliability analysis of the effect values. Second, the scholars may discover the shortcomings of the individual researches via the Meta-analysis, thereby avoiding the flaws. Third, the results of individual studies are aggregated through the Meta-analysis; hence, more objective explanations are conducted, and more reliable and accurate conclusions are researched. Fourth, new research directions and suggestions will be put forward after the Meta-analysis, which will provide references for further researches.

Nevertheless, like other methods, there are some flaws in the Meta-analysis. First, it takes a long period of time to conduct the Meta-analysis, due to its requirement for a large amount of literature. Thus, the workload is heavy. Second, it is difficult to find all the literature on a particular issue; hence, researchers cannot conduct a comprehensive statistical analysis. Last, it is difficult to ensure that the thinking and structure of the existing studies are consistent, that the qualities of these essays are high, and that the data is complete. Therefore, not all the documents are useful, and the amount of the literature may vary significantly among different research groups.

2 Research Method

2.1 Literature Retrieval

Although creative performance is a relatively new topic, scholars at home and around the world have conducted numerous researches. The paper firstly utilizes the keyword retrieval, the main method of literature retrieval, to search articles on the creative performance, employee behavior, and the factors affecting the employees' creative performance. Secondly, according to the references of the retrieved essays, the paper searches the related documents in each database, including the doctoral dissertations and journal articles published after 2010. The literature retrieval is conducted at the library of Sichuan University. Both Chinese databases (such as the CNKI, Wan) and foreign databases (for instance, Science-Direct, and Springer Link) are used in this research paper. Manual search is not required, since electronic magazines are available in the information age, and the databases above provide all the materials needed in this study.

2.2 Inclusion and Exclusion Criteria for Literature

(1) Literature Inclusion Criteria

First, the literature should be published or unpublished between 2000 to 2015.

Second, the literature should be associated with creative performance.

Third, the data should be relatively complete, including the values of correlation coefficients, such as T, F, and P. The articles may not contain all these data, but they should provide sufficient figures that are needed in this research.

Fourth, the research object should be explicit, describing the sample size, age, and gender.

Fifth, the factors for creative performance should be quantitatively determined.

Sixth, the questions listed in the scales should be expressed clearly.

Table 1. An overview of the selected literature

Number	Author	Year	Influencing factor	Research object	Effective sample size
1	Zhang Jian et al.	2013	The employees' emotions, orientations, etc.	Employees	222
2	Zhang Jian et al.	2013	Time pressure, creative personality, etc.	Freshmen	60
3	Dong Li et al.	2012	The employees' emotions, orientations, etc.	Freshmen	273
4	Zhang Jian et al.	2010	Independent motivation, control motivation, etc.	Employees	150
5	Tang ChaoYing et al.	2012	Leadership style, team recognition, etc.	Researchers	209
6	Song Yahui et al.	2015	Passion for work	Employees	162
7	Gong Zhenxing et al.	2015	Feedback, goal and personality, etc.	Employees	159
8	Wang Haocheng et al.	2013	Time pressure	College students	300
9	Li Xiulin et al.	2012	The employees' emotions	College students	91
10	Zhang Jian	2003	Working environment, motivation, needs, etc.	Employees, graduates	459
11	Gu Rui et al.	2015	Leadership style, etc.	Health care workers	101
12	Pamela Tierney	2004	The employees' creativity, leadership support, working experience, job complexity, etc.	Managers	12
13	Bartol Kathryn	2010	Participation in the creative process, etc.	Employees	86
14	Kwok Leung	2014	Work motivation, etc.	Employees	112
15	Yu-Shan Chan	2015	Environmental factors	College students	160
16	Kani Ulge	2016	Creative thinking, etc.	College students	216
17	Dong gun An	2016	The employees' motivations and personality, etc.	College students	143
18	Jin Nam Choi	2004	Personal factors, organizational environment, etc.	College students	430
19	Jan Kratzer	2004	Team communication, team recognition, etc.	Employees	243
20	Eric Chong	2010	Individual factors, supervision, and working environment, etc.	Employees	350
21	Wasilu Suleiman	2013	Working attitude, etc.	Employees	70
22	L. Olsson, S	2012	The quantity of managers, leadership style, etc.	Managers, employees	137

(2) Literature Exclusion Criteria

First, the articles largely about the literature review of this topic should be excluded.

Second, repeated reports should be excluded.

Third, the literature without relevant correlation coefficients or data should be excluded.

Fourth, the literature without explicit scale measurement questions should be excluded, since these articles cannot appropriately describe the research objects.

(3) An Overview of the Selected Literature

In the initial screening process of the Meta-analysis by using Stata, all literature adopts an empirical research with available data. According to different dependent variables, the selected articles are grouped into six categories: the employees' emotions (3 articles), time pressure (2 articles), creative personality (4 articles), work motivation (6 articles), leadership and teamwork (6 articles), and feedback consistency (2 articles). This paper examines which factor affects the employees' creative performance; hence, there is no causal relationship between these documents. In other words, the data in each article should be considered separately. After the searching and screening process, the selected articles are shown in Table 1: Considering the large data used in the Meta analysis via Stata (a software), it is inconvenient to list all of them in the table above. In the calculation process, however, the paper, according to a calculation method named the overall effect size, will select the complete and appropriate data from these research articles as the study data.

3 Research Results

3.1 Description of the Research Data

(1) Overall Effect Value

First of all, according to the above-mentioned method, the paper calculates and sorts the data extracted from the selected research articles. This study contains a total of nine research articles, each of which contain multiple influencing factors. Hence, a total of 25 effect values are obtained. Next, the paper presents a description of the selected research articles and effect values.

The selected articles with various effect values are divided into different types according to the reference sources. Table 2 illustrates the type and quantity of the research articles containing different effect values.

In Table 2, the "Journal" shows the effect value extracted from the journal articles, the largest proportion of which is 80%. "Thesis" indicates the effect value extracted from the doctoral dissertations and master's theses, which is 8%. The effect value extracted from the research reports is 12%. Figure 1 clearly illustrates these data.

Second, the paper calculates all effect values listed in the literature. This study mainly considers the factors that affect the creative performance of employees; therefore, all the determinants with data integrity mentioned in the

Table 2. The quantity of effect values from different article types

	Frequency	Percentage	Effective percentage
Journal	20	80%	80%
Thesis	2	8%	8%
Research report	3	12%	12%
Total	25	100%	100%

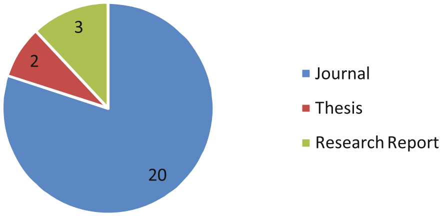


Fig. 1. Distribution of the quantity of effect values from different article types

selected documents are included in the table, even though they may influence each other. The specific data are shown in Table 3.

The distribution of the value of R and Z is clearly illustrated in Fig. 2, a line graph. In terms of the value of R, the average is 0.35; the standard deviation is 0.201; the maximum value is 0.71; the minimum value is -0.058; and the skewness is -0.219, which demonstrates negative skewness. As for the value of Z, the average is 0.62; the standard deviation is 0.410; the maximum value is 1.506; the minimum value is -0.086; and the skewness is 0.421, indicating positive skewness (See Table 4). Eventually, the paper, calculates the overall effect value of all literature, which is 0.551, generally indicating the positive correlation between the variables of the selected research articles and the employees' creative performance.

3.2 Homogeneity Test

The methods for homogeneity test include Q test, H test, chi-square test, forest plot. Using the Q test, the paper first calculates the Q value through the following equations:

$$Q = \sum\nolimits_{i=0}^{k} wi(ri - R)^2, \tag{1}$$

$$R = \sum wiri / \sum wi, \tag{2}$$

where k represents the number of the effect values of the selected literature; r is the effect value; and w denotes the proportion of the effect value. The calculated values of R and Q are 0.313575684 and 0.040220883, respectively.

Table 3. Related data for the selected literature

Number	Factor	R value	Z value	Sample size
1	Positive mood	0.228	0.301	495
2	Negative mood	0.054	0.082	495
3	Autonomous orientation	0.182	0.284	495
4	Control orientation	0.193	0.303	222
5	Features of time and pressure	0.33	0.543	60
6	Psychological satisfaction	0.419	0.718	60
7	Executive support	0.266	0.427	150
8	Need for autonomy	0.172	0.268	150
9	Capability requirement	0.243	0.387	150
10	Relationship requirement	0.277	0.447	150
11	Internal motivation	0.477	0.843	150
12	Motivation for identification	0.442	0.766	150
13	External motivation	−0.058	−0.086	150
14	Control motivation	−0.04	−0.06	150
15	Working motivation	0.395	0.669	150
16	Team performance	0.71	1.506	209
17	Team identity	0.509	0.917	209
18	Charismatic leadership	0.449	0.781	209
19	Concentration	0.585	1.11	162
20	Mastery approach goal orientation	0.485	0.861	162
21	Performance approach goal orientation	0.548	1.013	162
22	Performance avoidance goal orientation	0.348	0.577	162
23	Cognitive reevaluation	0.707	1.495	162
24	Inhibition of expression	0.431	0.743	162
25	Feedback consistency	0.33	0.543	159

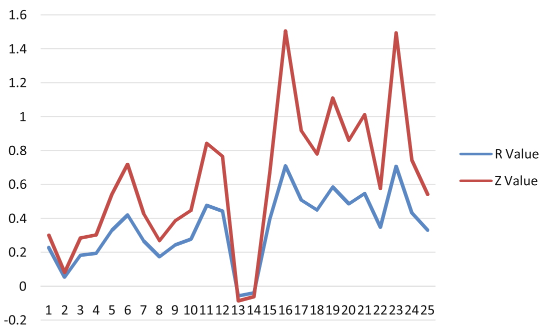


Fig. 2. Distribution of related data for the selected literature

Table 4. Description of statistical data of the selected literature

	N	Min	Max	Mean	Standard deviation	Skewness		Kurtosis	
							Standard error		Standard error
R value	25	0	1	0.35	0.201	−0.219	0.464	−0.158	0.902
Z value	25	0	2	0.62	0.41	0.421	0.464	0.184	0.902
Valid N (list state)	25								

The Meta results in Table 3 show that the 25 effect values are homogeneous. The Q test follows the chi-square distribution with a degree of freedom $K - 1$; hence, when the value of K is 25, the chi-squared value is 28.241 at a significance level of 0.05. In other words, if $Q = 0.040220883 < X^2_{0.02524} = 28.241$, then the selected research articles are homogeneous.

3.3 Bias Test

Bias test is considered essential in the meta-analysis. The publication bias refers to the different opportunities for publication and corresponding impacts on the results caused by the researchers, evaluators, editors' preferences of the direction and intensity of the research in the submission, acceptance, and publication process of an article. According to the book Conducting Meta-Analysis through Stata, there are multiple approaches to testing the publication bias, such as the funnel plot, Egger linear regression method, and Begg rank correlation method.

This paper uses the funnel plot, which is the most commonly used visualized method for a qualitative measurement of publication bias, first proposed by Light et al., in 1984. Figure 3 shows the funnel plot derived from the Meta-analysis of Table 3.

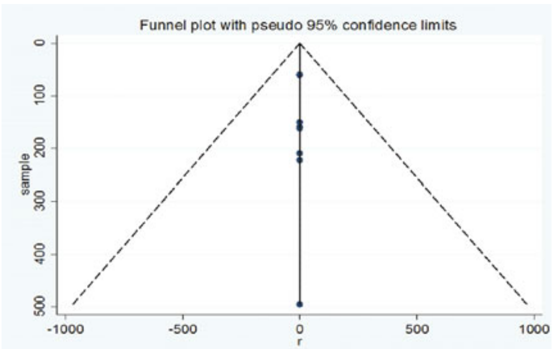


Fig. 3. The effect value of the selected literature

According to Fig. 3, there is no publication bias among the articles and theses used in this study, since the funnel plot is relatively symmetric.

4 Conclusions and Suggestions

It can be seen from overall effect values in Table 3 that when wholly considering the influencing factors (positive mood, negative mood, autonomous orientation, control orientation, time pressure, psychological satisfaction, supervisor support, autonomous need, relationship motive, internal motivation, external motivation, control motivation, working motivation, team performance, team identity, charismatic leadership, concentration, mastery approach goal orientation, performance approach goal orientation, performance avoidance goal orientation, cognitive reevaluation, inhibition of expression, feedback consistency), they are positively correlated to the employees' creative performance ($R = 0.551$). To enhance the creative performance, the managers may improve the employees' intrinsic motivation through incentive policies or through the executives' encouragement and support. Every employee, for the most part, will be affected by these factors; hence, this study, to some extent, may serve as a reference for the scholars or managers. Nevertheless, due to the possible interaction between these factors, their relationship with the employees' creative performance may be weak and accurate, which is the limitation of this paper. The author hopes that the future scholars of this field may clearly analyze this interaction.

We proposed suggestion as following: First, the future researchers may investigate the interaction between the factors impacting the employees' creative performance as well as explore the corresponding effects. The reason is because adjusting or changing one of the factors may indirectly influence other determinants, which contributes to a great change in the creative performance of the employees. Meanwhile, changing one factor may eliminate the effect of several factors, which may cause negative change in the staff creative performance. Second, the future researchers may examine the proportion of the effect of the individual factor on the creative performance. It assists managers to prioritize incentives so that the enterprise may maximize the profits at the lowest cost. Overall, joint efforts are needed in terms of enhancing the employees' creative performance.

Acknowledgements. This research was supported by Sichuan Social Science Planning Program Key Project (Grant No. SC16A006), System Science and Enterprise Development Research Center Project (Grant No. Xq16C12) and Sichuan University Central University Research Fund project (Grant No. skqy201653), Soft Science Project of Sichuan Science and Technology Department, (No. 2017ZR0033) and Sichuan Social Sciences Research 13th Five-Year annual planning project, (NO. SC16A006).

References

1. Chong E, Ma X (2010) The influence of individual factors, supervision and work environment on creative self-efficacy. Creativity Innov Manag 19(3):233–247
2. Dong L, Zhang J, Li X (2012) Relationship between employee's affect and creative performance under the moderation of autonomy orientation standard (in Chinese)
3. Gong Z, Zhang J, Zheng X (2013) Study on feedback-seeking influence of creative performance and motive mechanism (in Chinese)
4. Huang L, Krasikova DV, Liu D (2016) I can do it, so can you: the role of leader creative self-efficacy in facilitating follower creativity. Organ Behav Hum Decis Process 132:49–62
5. Song Y, He L et al (2015) Mediate mechanism between work passion and creative performance of employees. J Zhejiang Univ (Sci Ed) 42(6):652–659 (in Chinese)
6. Tang C, Liu Y, Wang T (2012) An empirical study of the relationship among charismatic leadership style, team identity and team creative performance in scientific teams. Sci Sci Manag S&T 33(10):155–162 (in Chinese)
7. Ulger K (2016) The creative training in the visual arts education. Thinking Skills Creativity 19(32):73–87
8. Zhang J (2003) Study on situational factors of employees creative performance and motive mechanism. Master's thesis, Capital Normal University, Beijing (in Chinese)
9. Zhang J, Dong L, Tian Y (2010) Improve or hinderinfluence of affect on creativity at work. Adv Psychol Sci 6:955–962 (in Chinese)
10. Zhang J, Song Y et al (2013) Relationship between employees' affect and creative performance the moderation role of general causality orientation. Forecasting 32(5):8–14 (in Chinese)

The Impact of Mixed Ownership Reform on Enterprise Performance–An Empirical Study Based on A-Share Listing Corporation in China

Sheng Ma(✉), Wenjie Li, and Sicheng Yan

Institute of Economics and Management,
Chengdu College, Chengdu, People's Republic of China
478767236@qq.com

Abstract. This paper, with data of 486 listed companies from 2003 to 2014 over the time span of 12 years, finds that the proportion of non state-owned shares in the Fixed-Effect Model is negatively correlated with the profitability of the company, and it is positively correlated with the company's development capacity; The proportion of circulating stocks has a significant positive impact on the corporate debt solvency and development ability; The proportion of the largest shareholder's shareholding and the proportion of the top ten shareholders are significantly correlated with the total asset growth rate, while separation rate of two rights and total asset growth rate are negatively correlated. Moreover, by using the Difference-in-Difference Model to test the reform policies, the regression results find that the restructuring has a significantly positive impact on corporate profitability and development capacity. The consistency of two methods confirms that the mixed ownership reform policy has a positive role in promoting the profitability, debt solvency and development capability of state-owned listed companies.

Keywords: Mixed ownership reform · Non state-owned shares · Enterprise performance

1 Introduction

The core reforms of state-owned enterprise (SOE) on the three levels are mainly: state-owned assets supervision, state-owned enterprise equity, state-owned enterprise operation. Among the three, SOE equity is the core of state-owned enterprise restructuring. The main core of SOE equity is to actively develop a mixed ownership economy, allowing more state-owned economy and other ownership economy to develop a mixed ownership economy, allowing mixed ownership economy to implement employee holdings, and finally the formation of co-ownership of capital and labor Therefore, the exploration of the relationship between the different ownership structure of various property nature and the performance of enterprises in the first restructuring process and the "second revision" is a

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_55

great guidance for the reform of SOE, and is conducive to a comprehensive understanding of the impact of restructuring on corporate performance.

In theory, the restructuring of SOE will improve the ownership structure of listed companies, improve the governance structure of listed companies, and then is conducive to improving the performance of SOE. This paper attempts to explore the impact of mixed ownership reform on corporate performance and analyze the policy implications of restructuring policy-making in the state-owned enterprises and non-state-owned enterprises, and, through theoretical and empirical analysis, tries to have some reference value to the participants and researchers of SOE reform.

2 Literature Review

2.1 Summary of Relevant Literature Abroad

Researches of foreign scholars on corporate restructuring in the background of the privatization began in the 1980s, and later expanded to a number of countries, including the restructuring of China's SOE [3,5,6]. Djankov and Murrell use the regression method of econometrics to make an empirical research on the enterprises in the transition economy and carry on the detailed model analysis [2]. Megginson and Nash studied the political, public institutions and economic factors in the state-owned enterprises privatization by shares issuance and sale on the public capital markets, conducted a comprehensive discussion of private ownership and supported the view that national policy and legal environment impact financial decisions [9]. Many foreign scholars have make related researches on the economic policy of corporate governance, such as Pagano and Volpin [10].

2.2 Summary of Relevant Literature in China

Since reform and opening up, China has been experiencing a slow process of economic transition. From the mid and late 1990s, due to the establishment of the modern enterprise system emphasized by the Third Plenary Session of the 14th CPC, the inefficiency of SOE drew the attention of the majority of Chinese scholars to the researches on the efficiency of SOE. Yao [15] based on the data analysis of the third industrial census, concluded that the rise of non-state economic components may promote the overall level of China's industrial improvement. Liu using the 1995 national industrial census data, found that among the variety of ownership enterprises, SOE has the lowest efficiency [7]. Liu and Li used [8] the data from 451 competitive enterprises (1994–1999) to make an empirical research and obtain the negative impact of state ownership on corporate performance, and conclude that non-state capital is positively correlated with firm performance. During this period, many scholars have studied the impact of ownership changes on corporate performance [4,13]. Song Ligang, Yao [12] also found a clear time trend, the effect of restructuring is most stable for those enterprises with a moderate length of restructuring history and for the enterprises that implemented restructuring between 1997 and 1999. Bai and Tao [1] used the panel data of the national industrial enterprises from 1998 to 2003 to analyze the economic and social benefits of the SOE restructuring, and concluded that the economic benefits of the state-controlled enterprises are better while the social benefits of the non-state-owned enterprises are better, and that the restructuring effect will continue for some time. Yang et al. [14] found that although the number of employees decreased after the restructuring of collective enterprises, the wage benefits per capita and taxation increased significantly, indicating that in general the restructuring of collective enterprises had a positive effect on social welfare. Sheng [11] used the multiplier method based on the tendency scoring matching to analyze the micro-data of Chinese industrial enterprises from 1999 to 2007, and concludes that the marketization and the introduction of competition mechanism makes the reform of SOE play a role in promoting social welfare. In the ongoing process of enterprise restructuring, some scholars use the Double-differential Model to compare the effects of the restructuring policy. Li and Qiao adopted the Double-differential Model for China's industrial data from 1999 to 2006, and found that the economic performance of state-owned enterprises improved significantly in 2003, and that the overall economic performance of SOE improved. Chen and Tang based on the national industrial data from 1999 to 2007, study the social burden and the policy burden on the enterprises with the Double-differential Model, and find that the mixed ownership reform can reduce the policy burden of SOR, and that the reform efficiency of monopoly mixed ownership is higher than that of the competitive industry.

3 An Empirical Study on Enterprise Performance by Mixed Ownership Reform

3.1 Data Sources

This paper uses the 486 listed companies listed in China's A-share market from 2003 to 2014, including all the industry classification of the SFC in 2014, with a total of 5832 effective observations, and set the dummy variable for the industry and the three economic zones. According to the "State Council's work report on the state-owned enterprise reform and development" published in 2012, more than 90% of the state-owned enterprises have completed the shareholding reform.

3.2 Definitions of Variables

(1) Corporate Performance Indicators

We use the company's financial indicators to measure the corporate performance indicators. ① Profitability is replaced by ROA of total assets. The return of the total assets(ROA) is the ratio of the net profit to the total assets of the enterprise over a certain period of time. The higher the value of ROA, the higher the profitability of the enterprise. ② Debt solvency is measured by the

Debt Asset ratio (DAR), which is the percentage of total liabilities at the end of the period divided by the total amount of assets. It is an important measure of the level of corporate liabilities and the degree of risk. ③ The growth capacity is measured by the sustainable growth rate (SGR) and the total asset growth rate (TAGR). The SGR is the highest growth rate that can be achieved by the non-issuance of new shares and maintaining the current operating efficiency and financial policy. So, this indicator represents a suitable pace of development The growth rate of the total assets is the ratio of the total assets increase of the year to its total assets at the beginning of the year, so it reflects the growth of the assets. The growth of the assets is an important aspect of the development of the enterprises. Enterprises of high growth rate can maintain the steady growth of assets.

(2) Explain the Variables

① Mixed ownership restructuring index: the proportion of non-state-owned shares (nonNSOS). Because the GTA CSMAR database contain the capital structure of all the listed companies including the total number of shares and the number of state-owned shares, so the proportion of non-state shares can be calculated.

② The proportion of circulating capital shares (LTBL): the ratio of the circulating shares to the total shares of the company. The greater the proportion of circulating shares, the more the stock reflects the true value of the company.

③ Proportion of sponsor shares (POP) is the ratio of the total number of sponsor shares to the total capital shares of the company. The sponsor shares refer to the special shares offered by the listed company to the founder(s) of the company.

④ The proportion of the largest shareholder holdings (POFLS): the ratio of the number of shares held by the largest shareholder to the total number of shares.

⑤ The proportion of the top ten shareholders holding (POTTS): the ratio of the number of shares held by the top ten shareholders to the total number of shares.

⑥ The separation rate of two rights (SRTR): the difference between the control of the actual controller of the listed company and his or her ownership.

We use the Herfindahl index to measure the concentration of the industrial market. The higher the Herfindahl index, the higher the degree of market concentration, and the higher the degree of monopoly. Other explanatory variables include total assets, total liabilities, paid-in capital, gross operating income, total operating costs, equity multiplier, equity ratio, flow ratio, quick ratio, total number of shareholders, and income tax.

As the listed companies that we have selected basically involve in the SFC industry classification of 2014, and these listed companies registered in the three major economic zones of the eastern and central regions. Therefore, for the study of the industry, several major industry categories are focused on, such as (1) electricity, heat, gas and water production and supply, (2) manufacturing, (3) real estate and (4) wholesale and retail. For the study of the regions, the three major economic zones in the eastern, central and western regions, two dummy variables are set as (East) for the East region and (West) for the West region. Of the 486 listed companies selected, 268 are located in the eastern region, 114 are in the central region and 106 are in the western region.

3.3 Design of the Measurement Model and Empirical Results

(1) The Impact of Restructuring on Corporate Performance of Listed Companies C An Analysis of Fixed-effect Model.

First of all, in order to study the impact of restructuring on the performance of listed companies, that is, the changes of the proportion of non-state equity in the overall equity changes, we need to control other factors that affect the performance. Both observable and unobservable factors can affect the corporate performance, so the observable factors are added to the regression model; for the unobservable factors that are not observable but will not vary over time, we can use the Fixed-Effect Model for their estimation, so the Fixed-effect Model is constructed as follows:

$$Y_{it} = \alpha + \beta_1 \text{nonNSOS}_{it} + \beta_i i = 26 X_{it} + \gamma_{it} \text{control}_{it} + \varepsilon_{it}. \tag{1}$$

Y_{it} is the performance index of the listed company i in the year t, LTBL_{it} is the proportion of circulating capital shares of the company i in the year t, and $\sum_{i=2}^{6} X_{it}$ are other important explanatory variables. The five explanatory variables are explained as follow: LTBL_{it} is the proportion of circulating capital shares of the company i in the year t, POP_{it} is the ratio of the total number of sponsor shares to the total capital shares of the company i in the year t, POFLS_{it} is the ratio of the number of shares held by the largest shareholder to the total number of shares of the company i in the year t, POTTS_{it} is the ratio of the number of shares held by the top ten shareholders to the total number of shares of the company i in the year t, and SRTR_{it} is the separation rate of two rights the company i in the year t. $\sum \text{control}_{it}$ is the other control variables used to improve the accuracy of the study. ε it is a random perturbation term and α, β_1, β_2, β_3, β_4, β_5, β_6 are the coefficients to be estimated. The results of model 1 are listed in Table 1.

From Table 1 on the results of restructuring effects, we find that: with the non-state-owned shares increasing, the profitability declines, sustainable growth rate decreases, but the total asset growth rate increases significantly. The proportion of circulating shares has a positive impact on the total assets growth rate and Debt Asset ratio. The proportion of circulating shares is correlated to the total asset growth rate at the significance level of 1%, and the proportion of circulating shares is correlated to Debt Asset ratio at the significance level of 5%.

(2) An Analysis of the Effectiveness of Mixed Ownership Reform Policy on Corporate Performance of Listed Companies C a Difference-In-Difference Model Analysis.

In order to further effectively separate the impact of mixed ownership reform on the performance of listed companies, the Difference-In-Difference can effectively overcome the selective deviation of the sample in the regression process.

Table 1. Impact of restructuring on corporate performance of listed companies (Fixed-effect Regression).

	Profitability Return of the total assets (ROA)	Debt solvency debt asset ratio (DAR)	Development capacity Sustainable growth rate (SGR)	Total asset growth rate (TAGR)
Proportion of Non state-owned shares	−0.100***	−0.12	−2.029***	0.217**
	(0.0374)	(0.0826)	(0.6470)	(0.0863)
Proportion of circulating shares	0.0109**	0.0363***	0.0643	0.0134
	(0.0048)	(0.0107)	(0.0836)	(0.0112)
Proportion of sponsor shares	0.0341**	−0.00746	0.912***	−0.0247
	(0.0170)	(0.0376)	(0.2950)	(0.0393)
Proportion of largest shareholder holdings	−0.000233	0.00125	0.00641	0.00367***
	(0.000517)	(0.00114)	(0.00895)	(0.00119)
Proportion of top ten shareholders holding	0.000611	0.00321**	0.00412	0.00381***
	(0.000574)	(0.00127)	(0.00992)	(0.00132)
The separation rate of two rights	−0.000218	−0.00216	−0.00559	-0.00433***
	(0.000684)	(0.00151)	(0.0118)	(0.00158)
The Herfindahl index	−0.0252	−0.352**	−0.247	−0.331**
	(0.0664)	(0.147)	(1.15)	(0.153)
Constant	−0.0659	0.924**	−1.417	−1.002**
Intercept term	(0.174)	(0.385)	(3.014)	(0.402)
Observations	4,111	4,111	4,110	4,107
R-squared	0.067	0.38	0.011	0.048

Note: The values in parentheses are the standard deviation. The superscripts ***, ** and * indicate that the estimated coefficients are significant at the 1%, 5% and 10% levels respectively.

Measuring and comparing the impact of restructuring policies on listed companies requires a more definite restructuring time point. From the micro level, we, in the process of handling the enterprise sample data, have recognized that China's restructuring process is a gradual reform, will always go through a number of steps or stages in order to achieve the final restructuring goals So the restructuring time point remains difficult to define from the micro level. At the macro level, the reform of state-owned enterprises has gone through five stages The emphasis on deepening the reform and on the development of mixed ownership economy began in the fourth stage with the "Decisions of the CPC Central Committee on Several Issues Concerning the Improvement of the Socialist Market Economic System" reviewed and approved by the Third Plenary Session of the 16th CPC Central Committee. Then the State Council SASAC has put forward two notices: "General Office of the State Council forwarded the SASAC Notice on Regulating the Reform of State-owned Enterprises" and "Opinions on Further Standardizing the Reform of State-owned Enterprises". Most of the equity nature changes of the listed companies from the CSMAR database occurred around 2006, and the number of publications on the restructuring of China's state-owned enterprises at the China National Knowledge Infrastructure (CNKI) reached its peak point in 2006. So in our literature the restructuring time point is roughly determined on this basis. Since the targets of SOE reform process are state-owned enterprises, and non-state enterprises are not involved in, so Difference-In-Difference model can be built. The specific procedures are as follow:

The data sample is divided into two groups: one is the state-owned enterprises labeled as $D1_i = 1$ (known as the test group, i.e., enterprise i is a state-owned enterprise), and the other group is non-state-owned enterprises (including private and foreign capital) labeled as $D1_i = 0$ (Known as the control group, i.e., enterprise i is a non-state-owned enterprise). Therefore, in the implementation of SOE restructuring policy, it is clear that only the SOE as the test group are affected, so the impact of restructuring policy on SOE can be expressed as E $(Y|D1_i = 1)$, and the impact of restructuring policy on non-state-owned enterprises can be expressed as $\mathrm{E}(Y|D1_i = 0)$. So we can express the net effect of restructuring policy on the performance of SOE as:

$$E(Y|D1_i = 1) - E(Y|D1_i = 0). \tag{2}$$

In order to explore the dynamic impact of the restructuring process, we carry out the time difference on this basis. Because the restructuring policy at different times will have differences, the test group and the control group will change with the policy time. One period labeled as $D2_i = 1$ is after the implementation of the restructuring policy, the other period labeled as $D2_i = 0$ is before the implementation of the restructuring policy. Therefore, we must also compare the performance changes before and after the restructuring policy, in which the performance impact after the implementation of restructuring policy can be expressed as $E(Y|D2_i = 1)$, and the performance impact before the implementation of restructuring policy can be expressed as $E(Y|D2_i = 0)$, so we can get the impact of restructuring policy at the time level:

$$E(Y|D2_i = 1) - E(Y|D2_i = 0). \tag{3}$$

So in order to measure the policy impact changes of the test group and the control group at the same time level, we can make the following adjustment to get the difference caused by restructuring policies:

Table 2. An analysis of the effect of mixed ownership reform policy on the performance of listed companies (difference-in-difference model regression).

	Profitability	Debt solvency	Development capacity	Total asset growth rate (TAGR)
	Return of the total assets (ROA)	Debt Asset ratio (DAR)	Sustainable growth rate (SGR)	
D1 before and after restructuring	0.00598	0.0269	−0.397*	0.141***
	−0.0139	−0.0307	−0.241	−0.0323
D2 test groupor	−0.0460***	0.00209	0.091	−0.0201
Control group	(0.0177)	(0.0392)	(0.307)	(0.0412)
D3 Interactive item	0.0566***	−0.0345	−0.0455	0.0852*
(D1*D2) coefficient of Policy effect	(0.0206)	(0.0455)	(0.357)	(0.0479)
nonNSOS	−0.095***	−0.125	−1.893***	0.171**
Portion of Non state-owned shares	(0.0371)	(0.082)	(0.643)	(0.0863)
LTBL	0.0110**	0.0371***	0.0591	0.0166
Proportion of circulating shares	(0.00476)	(0.0105)	(0.0826)	(0.0111)
POP	0.0309*	−0.00356	0.871***	−0.025
Proportion of sponsor shares	(0.017)	(0.0376)	(0.295)	(0.0396)
POFLS	−0.00029	0.000958	0.00697	0.00337***
Proportion of largest shareholder holdings	(0.000513)	(0.00113)	(0.00889)	(0.00119)
POTTS	0.000712	0.00313**	0.00499	0.00395***
Proportion of top ten shareholders holding	(0.000572)	(0.00127)	(0.00992)	(0.00133)
SRTR	−0.000392	−0.00223	−0.0056	−0.0043***
The separation rate of two rights	(0.000681)	(0.00151)	(0.0118)	(0.00158)
h	−0.0281	−0.295**	−0.5	−0.286*
The Herfindahl index	(0.0656)	(0.145)	(1.139)	(0.153)
Constant	−0.0407	0.727**	−0.864	−1.215***
	(0.167)	(0.369)	(2.889)	(0.388)
Observations	4,111	4,111	4,110	4,107
R-squared goodness of fit	0.7013	0.3809	0.0339	0.1518

Note: The values in parentheses are the standard deviation. The superscripts ***, ** and * indicate that the estimated coefficients are significant at the 1%, 5% and 10% levels respectively.

$$\{E(Y|D1_i = 1) - E(Y|D1_i = 0)\} - \{E(Y|D2_i = 1) - E(Y|D2_i = 0)\}. \quad (4)$$

The net effect of this restructuring policy not only measures the impact of the policy before and after its implementation, but also measures the policy differences between the test group and control group. So we have this regression model of Difference-In-Difference (Model 2):

$$Y_{it} = \alpha + \beta_1 D1_i + \beta_2 D2_i + \beta_3 D3_i + \beta_i i = 49X_{it} + \gamma_{it}\text{control}_{it} + \varepsilon_{it}. \quad (5)$$

Y_{it} is the performance index of the listed company i in the year t, and $D1_i$ is the dummy variable between groups, where $D1_i = 1$ is the test group, $D2_i$ is the time dummy variable; and $D3_i$ is the interactive item. $D3_i = D1_i \times D2_i$; β_3 is the Difference-in-Difference statistics, that is, the differences brought about by policy. Since other variables in Model Two are derived from the Model One, so the definition and interpretation of these variables in Model Two are not repeated here. For brevity we only list the regression results of several variables that related to the proportion of non-state equity, and the results and discussion of other variables are omitted. The results of Model (2) are shown in Table 2:

Table 2 shows the regression results for Model (2). The regression results between the companies that implement mixed ownership reforms and those that do not implement mixed ownership reforms show that difference-in-difference statistics for the Interactive item D3 is very significant for the total net profit margin, and is relatively significant for the total assets growth rate. In addition, the difference-in-difference statistics β_3 coefficient is significantly positive, indicating that compared to the enterprise without reform, the enterprise with reforms have greatly improved their profitability and development capability. The effect of the mixed ownership restructuring policy is not significant to the difference-in-difference statistics β_3 coefficient in the regression of the debt solvency, which indicates that the effect of the mixed ownership reform policy is not reflected. From Table 2, we can also see that with the increase in the proportion of non-state-owned shares, the decline in profitability and sustainable growth rate is relatively alleviated compared with the results in Table 1, indicating that the restructuring policy has played a role in improving the profitability of enterprises.

4 Conclusions and Suggestions

The empirical conclusion is as follows:

(1) With the increase in the proportion of non-state-owned shares, the total asset net profit margin decreases, and the sustainable growth rate decreases, but the total assets growth rate increases significantly; the proportion of circulating shares has a positive impact on the total assets growth rate and Debt Asset ratio; Both the proportion of the largest shareholders holding and the proportion of the top ten shareholders holding are positively correlated with the total asset growth rate, while the separation rate of two rights is negatively correlated with the total assets growth rate. The vertical comparison indicates that the increase

in the proportion of non-state equity can improve the development capacity of enterprises, increase its debt solvency and reduce its profitability.

(2) The policy effect of Mixed ownership reform: Compared with enterprises that have not been reformed, the enterprises that are reformed have a greater improvement in profitability and development capability. The policy effect of mixed ownership reform is not reflected in the debt solvency the results of other explanatory variables in the Difference-in-difference Model are in good agreement with the regression results of the Fixed-effect Model, which indirectly confirm that the mixed ownership reform policy has a positive role in promoting these financial measure of state-owned listed companies such as profitability, debt solvency and development capability.

References

1. Bai D, Lu J, Tao Z (2006) An empirical study on the effect of state-owned enterprise reform. Econ Res 8:4–13 (in Chinese)
2. Djankov S, Murrell P (2002) Enterprise restructuring in transition: a quantitative survey. J Econ Lit 40(3):739–792
3. Frydman R, Gray C, Hessel M et al (1998) When does privatization work? The impact of private ownership on corporate performance in the transition economies. Q J Econ 114(4):1153–1191
4. Hu Y, Song M, Zhang J (2005) The relative importance and interrelationship of the three theories of competition, property right and corporate governance. Econ Res 9:44–57 (in Chinese)
5. Kang SW, Lee KH (2016) Mainstreaming corporate environmental strategy in management research. Benchmarking 23(3):618–650
6. Kim J, Song HJ et al (2017) The impact of four CSR dimensions on a gaming company's image and customers' revisit intentions. Int J Hospitality Manage 61:73–81
7. Liu X (2000) The impact of the ownership structure of china's industrial enterprises on the efficiency difference - an empirical analysis of the census data of national industrial enterprises in 1995. Econ Res 2:17–25 (in Chinese)
8. Liu X, Li L (2005) An empirical analysis of the impact of restructuring on corporate performance. China Ind Econ 3:5–12 (in Chinese)
9. Megginson WL, Nash RC, Netter JM et al (2004) The choice of private versus public capital markets: evidence from privatizations. J Finan 59(6):2835–2870
10. Pagano M (2005) The political economy of corporate governance. CSEF Working Pap 95(4):1005–1030
11. Sheng D (2013) The restructuring of state-owned enterprises, the degree of competition and social welfare - a study based on the corporate cost mark-up percentage. Economics (Q) 4:1465–1490 (in Chinese)
12. Song L, Yao Y (2005) The impact of restructuring on corporate performance. Soc Sci China 2:17–31 (in Chinese)
13. Wu C, Li D (2005) Enterprise behavior in mixed market. Donyue Tribune 1:38–47 (in Chinese)
14. Yang Z, Lu J, Tao Z (2007) Political asylum and reform: research on the restructuring China's collective enterprise. Econ Res 5:104–114 (in Chinese)
15. Yao Y (1998) Effects of non-state ownership on technological efficiency of industrial enterprises in China. Econ Res 12:29–35 (in Chinese)

Online Impulse Buying on "Double Eleven" Shopping Festival: An Empirical Investigation of Utilitarian and Hedonic Motivations

Umair Akram[1], Peng Hui[1], Muhammad Kaleem Khan[1], Muhammad Hashim[2(✉)], Yixin Qiu[1], and Ying Zhang[1]

[1] School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
[2] Faculty of Management Sciences, National Textile University, Faisalabad 37610, Pakistan
hashimscu@gmail.com

Abstract. "Double Eleven" online shopping festival has become the most celebrated and largest online shopping event in China. This phenomenon originates the need to explore impulse buying behavior in relationship of double eleven festival. This study aims to investigate the structural relationship between hedonic, utilitarian web browsing and online impulse buying behavior on "Double Eleven" shopping festival. Data was collected from 426 online shoppers in six different districts of Beijing. In order to assess the measurement model and to test the hypothesis structural equation modeling was utilized by using AMOS 21. Additionally, Reliability, Discriminant and convergent validity was used for the results of proposed model. The results of this study found that utilitarian and hedonic web browsing positively influence on online impulse shoppers with regard to "Double Eleven" shopping festival in China. To the best of our knowledge, this study is pioneer, aiming at exploring the dynamics of online impulse buying behavior with utilitarian and hedonic predictors with regard to biggest online shopping festival in China. This study provides a deep understanding of e-shoppers toward online impulse behavior in the field of Chinese e-commerce industry that can be generalized to other countries also. Implications for scholars and e-tail managers of our study are discussed.

Keywords: E-commerce · "Double Eleven" shopping festival · Online impulse buying · Utilitarian web browsing · Hedonic web browsing

1 Introduction

With the growing field of e-commerce, online shopping trend is becoming popular across the world. China has become largest e-commerce economy in the world. Consumers always prefer to buy goods online especially on occasions. Most of

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_56

the time when they make decision about online shopping they act impulsively. One click ordering, easy access to products, absence of delivery, rich information about product and time saving these all factors leads to buy impulsively online. There are many factors that influence on impulse buying behavior. Impulse purchase, is explained as compelling, hedonically complex and unplanned buying behavior [35]. With the tremendous growth of e-commerce and rapid development of information technology, online impulse purchasing has become an epidemic. It is also noteworthy that about 40% of all online shoppers spending is attributable to online impulse purchasing [24,36]. Online shopping setting is more encouraging to impulse buying behavior than its offline complement. Online shopping environment liberates consumers from the constraints (e.g., social pressure from the staff, inconvenient store locations, limited operating hours, and it need much time) that they might experience during physical shopping events [13].

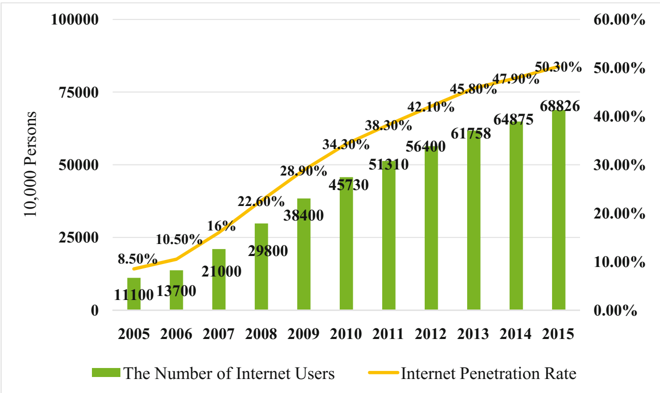


Fig. 1. The size of Chinese internet users and penetration rate

According to the CNNIC report in December 2015 China had 688 million netizens, up 39.51 million over the previous year. The internet penetration rate reached 50.3%, up 2.4% points from the end of 2014 (see Fig. 1) [8]. The survey result shows that in 2015, the mobile phone was the most popular device used for internet access by new internet users, by 71.5% of them, up by 7.4% points from the end to 2014. Among the new internet users in 2015, 46.1% were under 19, and 46.4% were students, and entertainment and communication were their two biggest reasons to access the internet. Online impulse purchase is defined as an immediate and sudden online buying with no preshopping intentions [31]. Previous studies reported that unplanned buying account for up to 60% of all purchases [19,27] and according to [18,21] 40% to 60% of impulse purchases depending on product category. A number of studies have been done in offline setting with impulse buying while a little attention focused on online impulse buying. Recently in an online buying setting, scholars have studied how to better appeal to impulse shoppers to take benefit of the behavior which has helped brick and mortar retailers flourish for decades [9,23,36,37]. Irrespective of context, a main purpose in retailing is to increase the attraction to improve sales [6,21]. Due to the pervasiveness and practical implications of impulse purchase, retailers have focused significant efforts on facilitating the behavior [9,23]. This study is not only beneficial for online retailers, but it also has future directions and guidelines for scholars who are having been studied on impulse buying behavior and produced number of studies in last decade.

In this regards, it is important to understand the comparative benefits of buying online over offline shopping. Therefore, retailers should be aware of the motivations behind the consumers' online buying process. Motivations are an essential factor to define the individual's behavior. They stem unmet needs and represent, through concrete actions, the benefits which people hope to achieve [33]. Two types of motivational factors influence on online shopping i.e., Utilitarian and Hedonic motivations [26]. While making essential consumption decisions, it is reasonable to create a difference between utilitarian and hedonic motives. These motives construct two aspects of attitudes with respect to behavioral acts. Hedonic motivations are related with experiential or emotional aspects which make the pleasured shopping experience while utilitarian motivation aspects associated with rational, functional, practical, economic or extrinsic benefits. In this research, we focus on how these two motivations influence on online impulse shopping.

Table 1. 2009–2016 Alipay "Double Eleven" day trading volume (unit: 100 million yuan)

Year	2009	2010	2011	2012	2013	2014	2015
Turnover	0.5	2.5	53	181	350.19	571.12	912.17

Source: Taobao official website.

In China, online shopping market has shown great potential due to rapid growth of Chinese netizens and online consumers (see Fig. 1). In order to achieve the competitive advantages a number of businesses joining the e-commerce industry. They must understand consumer needs and try to meet their demands, they should make a marketing strategies (e.g., low price product, augmented services, speedy transaction process, quality of products) to earn profit and attract netizens. In this regards, the "Double Eleven" online shopping festival has an emerging trend. Taobao introduced "Double Eleven" shopping festival on November 11, 2009 on "singles' Day". "Double Eleven" online shopping festival is most popular among Chinese netizens/e-commerce industry and it is largest commercial activity. Early 2016, the festival has been successfully held for seven times. This event has carried out vast economic benefits, which have increased year by year, and every year revive the record single-day net buying business amount (see Table 1). Aforesaid studies about online impulse buying behavior have shown that there is a great opportunity to identify how online impulse purchase effected due to this event i.e. "Double Eleven" because authors best knowledge no study has been conducted in this context. Only a few researchers have studied to examine the "Double Eleven" model successful factors from different perspectives. Liu [24] reported that this setting can effectively motivate to online shoppers to create impulse buying behavior (buying behavior that goes beyond what has been planned).

This study provides a better understanding about motivational factors that influence on online impulse buying adoption in China with the perspective of "Double Eleven" shopping festival. A little work has been investigated on relationship between motivational factors and online impulse buying, but no single study has yet investigated these factors in the context of "Double Eleven" shopping event. The Current study fill this gap by incorporated the conception of motivational factors i.e. Utilitarian and Hedonic on online impulse buying behavior in the context of "Double Eleven" shopping activity. Aforementioned discussions, the purpose of this manuscript is threefold: (1) to highlights the major findings in study regarding "Double Eleven" shopping festival, (2) to investigate the relationship between online impulse buying and utilitarian motivations with the regard of Double Eleven shopping activity (3) to explore the relationship between hedonic motivations and online impulse buying with the regard to "Double Eleven" shopping activity. The rest of the manuscript is organized as follows. First we discuss core concepts of online impulse buying and motivational factors i.e., utilitarian and hedonic and their relationship, based on their connection we develop the hypotheses and draw research model. The next section, we test the hypotheses based on data collection and data analysis. Finally, last section presents the research outcomes, managerial implications and future direction.

2 Online Impulse Buying

Online impulse purchasing is an immediate, unplanned, compelling and sudden purchase behavior while shopping experience [1,2]. Consumers usually tend to make immediate and unintended purchases while making online buying [20] their intention about online buying might be related to website complicity and simplicity [38]. In this context, Sharma [34] demonstrated that online buying is driven by low cognitive control or spontaneous behavior and consumers' emotions. They argued that consumer's impulse buying behavior is driven by appealing objects, which influence them to make purchase without considering financial and other factors of the online buying. Verhagen et al. [36] and Dholakia [9] mentioned that online consumers are more likely to buy impulsive as compared to offline. Many consumers to overspend due to online transactions because the virtual procedure does not feel like spending money [11]. Online marketing stimuli make buying spontaneously and allow online buyers to be less risk-averse [25]. E-impulse influenced by many motivational factors i.e. utilitarian and hedonic web browsing and other factors such as web site quality, sales promotion, ease of payment.

3 Utilitarian and Hedonic Motivations

According to the scholars [25,28] Utilitarian and hedonic factor both influence on impulse purchase. Impulse purchasing and different kind of browsing are effortless feelings [34]. Online shoppers motivate by searching due to the benefits such as entertainment, uniqueness and fun [15]. For specific products, online impulse buying occurs due to emotional and hedonic browsing [30]. Internet facilitates browsing the online merchandise for informational purpose (i.e., utilitarian) and recreational i.e., hedonic browsing [25]. Consumers usually act as impulsively when they are taking online decisions due to easy buying e.g., one click order, absence of delivery efforts and easy access to product [36]. Park [9] confirmed that utilitarian and hedonic browsing has a relationship with impulsiveness in the context of online buying of apparel products. Additionally, Gohary [14] demonstrated that effect of hedonic and utilitarian browsing on impulse behavior which points towards the importance of hedonic and utilitarian browsing for impulse purchasing over the internet [12,36]. Finally, Rezaei [32] confirmed that utilitarian and hedonic browsing positively influence on online impulse buying. Based on above studies the following hypotheses regarding the relationship between motivational factors and online impulse buying is derived:

H1 (HWB → OIB): Hedonic web browsing positively relate with online impulse buying in the context of "Double Eleven" shopping events.
H2 (UWB → OIB): Utilitarian web browsing positively relate with online impulse buying in the context of "Double Eleven" shopping events.

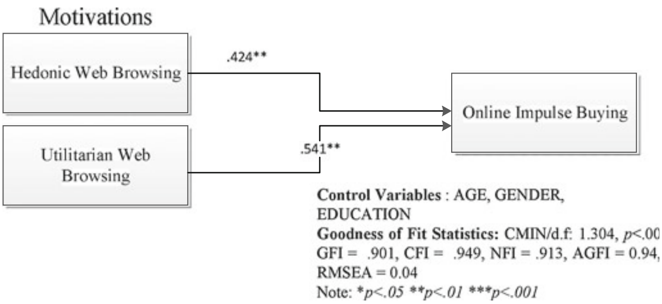


Fig. 2. A structure model for Hedonic web browsing, Utilitarian Web browsing and Online impulse buying

Mathematical Model

$$\text{OIB} = \beta_0 + \beta_1\,(\text{HWB}) + \beta_2\,(\text{UWB}) + \varepsilon.$$

4 Research Methodology

The aim of this study is to find the motivational factors effect on OIB on "Double Eleven" shopping festival. Data was collected from six different districts i.e. Haidian, Chaoyang, Fengtai, Changping, Tongzhou, and Dongcheng of Beijing. The survey was conducted on the Double Eleven shopping festival date on 11th November 2016. Convenience sampling method was utilized to collect the large data.

4.1 Questionnaire

Paper questionnaire and online survey techniques was used for data collection. Initially, questionnaire was developed/adapt in English and subsequently translated into Chinese and back translation into English for accuracy. Services of Chinese language experts and translators were employed to translate the questionnaire into Chinese and then it was also translated back into English to check accuracy. Translation and back translation was validated by International Chinese Training Center, Beijing. The first part of questionnaire was designed to ensure the respondents have online buying experience on "Double Eleven" shopping festival and in order to get information related with respondents' characteristics. The second section was designed to examine the relationship among online impulse buying behavior, hedonic and utilitarian motivations. Pilot testing were utilized ($n = 65$). Additionally, total 470 questionnaires were distributed among online shoppers and ($n = 426$) was valid for the data analysis with response rate 90.63%. Note books were provided to the respondents in order to motivate and seduce their patience, which cost about 5 RMB (.7 $).

4.2 Instrumentations

Five items were adapted from [36] to measure the online impulse buying behavior e.g., My purchases were spontaneous. In order to measure the utilitarian web browsing five items were adopted from [9] e.g., I browse to buy better items in price and quality and to measure the hedonic web browsing four items were adopted from [9] e.g., while web browsing, I am able to forget my problems and to feel relaxed. Rezaei [32] also used these three instruments in his study.

5 Results and Hypothesis Testing

Table 2 reports the demographic characteristics of the respondents. Total 470 questionnaires were floated, out of 470, 426 respondents were found who had online shopping experience. Demographics show that there are higher number of female respondents as compared to males i.e. 63% are females and 37% males. Maximum respondents were belonging to 25–31 years of age (i.e. 38.6% of total respondents). Final collected respondents came out to be 51.8% students, 20.6% employed, 11.2% individuals having managerial job and 9.8% having their own business.

Table 2. Respondents Demographic Information ($n = 426$)

Age		Gender		Education	
18–24	32.1%	Male	37%	Doctorate	11.5%
25–31	38.6%	Female	63%	Master	23.2%
32–40	25.3%			Bachelor	35.3%
above 40	4.0%			Diploma	25.0%
				Other	5.0%
Occupation		**Years using the internet**		**Online Impulse Experience**	
Business Owner	9.8%	1–4	7.5%	Yes	426
Managerial level	11.2%	4–9	53.1%	No	470
Employee level	20.6%	10–14	35.6%		
Student	51.8%	More than 14	3.8%		
Other	6.6%				

Kaiser-Meyer-Olkin (KMO) test formula

$$\mathrm{KMO}_j = \frac{\sum_{i\neq j} r_{ij}^2}{\sum_{i\neq j} r_{ij}^2 + \sum_{i\neq j} u_{ij}^2},$$

$$\mathrm{AVE} = \left(\sum_{i=1}^{k} \lambda_i\right)^2 \Bigg/ \left[\left(\sum_{i=1}^{k} \lambda_i\right)^2 + \sum_{i=1}^{k} (1-\lambda_i^2)\right].$$

Measured with the Joreskog Rho $= \sum_{i=1}^{k} \lambda_i^2 \Big/ \left[\sum_{i=1}^{k} \lambda_i^2 + \sum_{i=1}^{k} (1-\lambda_i^2)\right]$.

Before examining the structure model, multi-collinearity issue was assessed through variance inflation factor (VIF) produced by SPSS 21 version. The results of VIFs indicated the value of 1.34 for online impulse buying, 2.70 for utilitarian web browsing and 1.76 for hedonic web browsing variable, all stated values are lower than the threshold value 10 [10]. Therefore, the results show that no serious problem of multicollinearity in our study. Missing values in data set is a big challenge for social science scholars like information system, Human resource management and marketing areas. Many techniques can be performed for the treatment of missing values but multiple imputation considered effectively [32]. In this study, to handle missing values and impute missing values effectively by using SPSS, expectation maximization algorithm (EMA) was utilized. For the missing value treatment, little's missing completely at random MCAR which is chi-square χ^2 test for missing completely at random Little's MCAR test: $\chi^2 = 334.270$, $df = 332$, significance level = .444. Furthermore, before factor analysis Kaiser-Meyer-Olkin (KMO) test was assessed for the sampling adequacy which shows KMO value .867. KMO value .867 is less then benchmark value .10 [16].

To examine the structure model confirmatory factor analysis (CFA) was used to test the hypothesis and measurement model by using AMOS 21 version. After removing problematic items, we reran the CFA analysis, the results indicated that satisfactory fit (CMIN/DF: 1.82; $p < .000$; NFI = .98; IFI = 0.96;

Table 3. Convergent Validity and Reliability

Construct	Item	Loadings	Cronbach's alpha	Composite reliability[a]	AVE[b]
Online Impulse Buying (OIB)	OIB1	0.875	0.894	0.98	0.81
	OIB2	0.863			
	OIB3	0.924			
	OIB4	0.911			
	OIB5	0.889			
Utilitarian web browsing	UWB1	0.945	0.877	0.912	0.9
	UWB2	0.865			
	UWB3	0.876			
	UWB4	0.878			
Hedonic web browsing	HWB1	0.888	0.845	0.92	0.79
	HWB2	0.899			
	HWB3	0.924			
	HWB4	0.789			

Notes: [a]Composite reliability (CR) = (Square of the summation of the factor loading)/(square of the summation of the factor loadings) + (square of the summation of the error variance); [b]Average variance extracted (AVE) = (summation of the square of the factor loadings)/(summation of the square of the factor loadings) + (summation of the error variances).

Table 4. Results of structural relationships and hypothesis testing

Construct	Mean	Std. Deviation	OIB	UWB	HWB
OIB	4.03	1.56	0.81	-	-
UWB	3.98	1.77	0.54	0.9	-
HWB	4.16	1.55	0.452	0.64	0.79

Note: Acronyms: Online impulse buying (OIB); Utilitarian web browsing and hedonic web browsing; Diagonal bold values represent AVEs and off-diagonal denotes square correlation among constructs. [$*p < 0.05$, $**p < 0.01$, $***p < 0.001$].

GFI = .887; CFI = .91; RMSEA = 0.05). All values of CFA analysis were satisfactory and meet the threshold value. Reliability and Validity test performed in the model. Convergent validity was examined by using three parameters i.e., (1) factor loading grater then .70 with statistical significance (2) composite reliability (CR) larger then .80. (3) Average variance extraction (AVE) higher then .50 [16]. Table 3 indicated that all values of factor loading is greater than .70, in addition, all constructs show the high level of reliability internal consistency CR and Cronbach's alpha values ranges from .84 to .98. The value of AVE is greater then the benchmark value .50. Thus, we attained the good level of convergent validity and reliability. Table 3 reports all statistical values of mean, standard deviation and discriminant validity. We used Hair's criterion to evaluate the discriminant validity which directs square root to average variance extraction AVE for all constructs should be higher then it correlations [16]. Discriminant validity among all variables were confirmed based upon Hair's criterion (see Table 4).

Table 5. Results of structural relationships and hypothesis testing

Hypotheses	Path	R^2	Path Coefficient	Standard error	t-statistics	Decision
H1	HWB → OIB	0.424	0.823	0.067	21.4***	Supported
H2	HWB → OIB	0.541	0.764	0.089	32.1***	Supported

Notes: t-values for two tailed test: ***2.58 (significance level = 1%); **1.96 (significance level = 5%); and *t-value 1.56 (significance level 10%) [17]; Acronyms: online impulse buying (OIB); Hedonic web browsing (HWB) and utilitarian web browsing (UWB).

Table 5 depicts the results of path coefficients (β) which indicates the hypothesized association among the constructs. Figure 2 represents the results of structure model. *H1* that proposes there is a positive relationship between hedonic web browsing and online impulse buying was approved with path coefficient (β) 0.823, t-statistics 21.4 and standard error 0.067. Results demonstrated that hedonic web browsing positive and significantly influences on online impulse buying on "Double Eleven" shopping festival. *H2* proposed that utilitarian web browsing positively influence on online impulse buying (utilitarian web browsing → online impulse buying) with coefficient (β) of 0.764, standard error of 0.089 and the value of t-statistics 32.1. *H2* was also supported. Furthermore, Table 4 shows the values of R^2 for the relationships of constructs. The R^2 value for (UWB → OIB) .541 and for (HWB → OIB) .424, which means 54.1% change in online impulse buying due to utilitarian web browsing while 42.4% change in online impulse buying due to hedonic web browsing. Hedonic web browsing strongly influences on online impulse buying than utilitarian web browsing. The findings confirm that UWB and HWB strongly influenced on online impulse buying.

6 Discussions and Conclusion

With a remarkable growth of Chinese internet users and tremendous penetration rate of internet (see Fig. 1), web browsing motivation is considered as an essential part of online shoppers' experience. The current study provides important insights of online impulse behavior based on web browsing motivation i.e., Hedonic and Utilitarian for e-tailers to develop their e-business strategies specially on big event like "Double Eleven". This study investigates online impulse buying shoppers towards "Double Eleven" e-shopping festival and examines whether hedonic and utilitarian web browsing have any impact on it. This manuscript contributes to existing studies regarding online impulse buying behavior by using exogenous constructs i.e., hedonic and utilitarian web browsing. With the remarkable growth of online business, "Double Eleven" has turned out not only a festival to celebrate, but also most popular and highest sales volume festival in online shopping. This study provides a significant insight of utilitarian and hedonic web browsing behavior which influence on online impulse buying behavior. In China, e-commerce companies are facing high rivalry environment and they need to make innovative e-business strategies in order to increase online impulse buying, that makes the results of this paper useful.

The results of this study demonstrated that Hedonic and utilitarian web browsing both influenced on online impulse buying behavior on “Double Eleven” shopping festival. The findings also support broadened theory of impulse purchase behavior [5], which recommends that web browsing motivation is a key to enhance online impulse purchase for apparel purchase from both hedonic and utilitarian perspective. Inclination with the findings of this study Kim and Eastim argued that there is a huge difference between hedonic and web browsing behavior in online shopping. With the perspective of utilitarian value, consumers are focused towards completing consumption goals [3,4,7,29], while hedonic value, consumers are more attentive on entertainment, emotional and fun when dealing with online browsing [7,22]. Park [9] found that hedonic web browsing positively influences on OIB whereas utilitarian web browsing negatively influenced on online impulse buying. In an online shopping context, this evidence supports the hedonic nature of online impulse behavior [6,30].

E-tailer should make successful strategies by focusing on hedonic and utilitarian web browsing motives for e-shoppers by ensuring in professional manner, for instance; easy purchase process, good selection of variety, elegance and security of the web site. These all factors lead to online impulse shopper. This manuscript has some important practical implication towards OIB on “Double Eleven” shopping activity. This study’s finding will help the e-tail managers and web hosts associated with online selling to exploit exogenous factors HEB and UWB to enhance their impulse sales through web traffic. Utilitarian and hedonic web browsing plays a significant role to motivate e-shoppers to buy impulsively specially on “Double Eleven” event. In China, there is rapid growth towards netizens and penetration rate of internet users, now every individual prefer buying online. The outcomes of this study will useful for two e-tail big players in China (i.e., www.jd.com and www.taobao.com) as well as small e-tail businesses. Therefore, marketers should focus on these strategies in order to increase the impulse buying behavior which is major contribution in the field of e-commerce retailing and marketing (Table 6).

Despite all the strengths and potential contributions of the study, to generalize the findings, future research should consider the limitation of study. First, in this study we collect data from Beijing and just focus on “Double Eleven” shopping festival. Results can be enhanced by expanding the study sample and incorporating more festivals that are found to increase online sales. Moreover, this manuscript has considered only two motivational factors i.e., UWB and HWB, future scholars may use other factors related to web site personality, emotional factors such as perceived usefulness and perceived ease of use and outcomes could be used to improve other type of e-commerce web site. This study is based on quantitative data analysis and cross sectional in nature. For online shopping characteristic of Double Eleven festival longitudinal and qualitative method (i.e., interviews and focus group) may also be used for deep understanding of online shopping environment. Pre-and post-purchase behavior with regard to “Double Eleven” shopping festival may also be analyzed by future researchers. Lastly, future studies are recommended to analyze the proposed model in different e-tail cultural environment.

Acknowledgements. The authors want to extend their gratitude towards the Editor (Jiuping, Xu) and the anonymous Reviewers for their indispensable and valuable suggestions and comments that improved the quality of the paper significantly. Furthermore, they thank Prof. Peng Hui, Professor of finance, School of economics and management, Beijing University of Posts and Telecommunications for valuable suggestions.

Appendix

Table 6. Measurement items

No	Construct	Scale	Source
1	Utilitarian web browsing	UWB1 I browse to buy better items in price or quality	[30]
		UWB2 I browse the shopping web sites to gather information about products	
		UWB3 I look around the shopping web sites to comparison shop	
		UWB4 I browse the shopping web sites to get additional value as much as possible	
		UWB5 I browse for efficient shopping online	
2	Hedonic web browsing	HWB1 While web browsing, I am able to forget my problems and to feel relaxed	[30]
		HWB2 During web browsing, I am very excited, like playing	
		HWB3 I enjoy web browsing enough to forget a time out	
		HWB4 I look around at items on the internet just for fun	
3	Online impulse buying	OIB1 My purchase was spontaneous	[36]
		OIB2 My purchase was unplanned	
		OIB3 I did not intend to do this purchase before this shopping trip	
		OIB4 Before visiting the site, I did not have the intention to do this purchase	
		OIB5 I could not resist doing this purchase at the site	

References

1. Akram U, Peng H et al (2017) Impulsive buying: a qualitative investigation of the phenomenon. Springer, Singapore
2. Akram U, Peng H et al (2016) Impact of store atmosphere on impulse buying behaviour. Moderating effect of demographic variables. Int J U-and E-Serv Sci Technol 9(7):43–60

3. Babin BJ, Griffin M (1994) Work and/or fun: measuring hedonic and utilitarian shopping value. J Consum Res 20(4):644–656
4. Babin LA, Babin BJ, Boles JS (1999) The effects of consumer perceptions of the salesperson, product and dealer on purchase intentions. J Retail Consum Serv 6(6):91–97
5. Baumeister RF (2002) Yielding to temptation: self-control failure, impulsive purchasing, and consumer behavior. J Consum Res 28:670–676
6. Beatty SE, Ferrell ME (1998) Impulse buying: modeling its precursors. J Retail 74(2):161–167
7. Choi J, Li YJ et al (2014) The odd-ending price justification effect: the influence of price-endings on hedonic and utilitarian consumption. J Acad Mark Sci 42(5):545–557
8. CNNIC (2016) Statistic report on internet development in China
9. Dholakia UM (2000) Temptation and resistance: an integrated model of consumption impulse formation and enactment. Psychol Mark 17(11):955–982
10. Diamantopoulos A, Siguaw JA (2006) Formative versus reflective indicators in organizational measure development: a comparison and empirical illustration. Br J Manage 17(4):263–282
11. Dittmar H, Long K, Meek R (2004) Buying on the internet: gender differences in on-line and conventional buying motivations. Sex Roles 50(5):423–444
12. Eastin SK, Matthew S (2011) Hedonic tendencies and the online consumer: an investigation of the online shopping process. J Internet Commer 10(1):68–90
13. Eroglu SA, Machleit KA, Davis LM (2001) Atmospheric qualities of online retailing: a conceptual model and implications. J Bus Res 54(2):177–184
14. Gohary A, Hanzaee KH (2014) Personality traits as predictors of shopping motivations and behaviors: a canonical correlation analysis. Arab Econ Bus J 10(9):166–174
15. Ha S, Stoel L (2012) Online apparel retailing: roles of e-shopping quality and experiential e-shopping motives. J Serv Manage 23(2):197–215
16. Hair JF, Black WC et al (2010) Multivariate data analysis: a global perspective
17. Hair JF, Sarstedt M (2011) PLS-SEM: indeed a silver bullet. J Mark Theor Pract 19(2):139–151
18. Hausman A (2000) A multi-method investigation of consumer motivations in impulse buying behavior. J Consum Mark 17(5):403–426
19. Inman JJ, Ferraro R, Winer RS (2004) Where the rubber meets the road: a model of in-store consumer decision making
20. Jones MA, Reynolds KE et al (2003) The product-specific nature of impulse buying tendency. J Bus Res 56(7):505–511
21. Kacen JJ, Hess JD, Walker D (2012) Spontaneous selection: the influence of product and retailing factors on consumer impulse purchases. J Retail Consum Serv 19(6):578–588
22. Kaltcheva VD, Weitz BA (2013) When should a retailer create an exciting store environment. J Mark 70(1):107–118
23. Kervenoael RD, Aykac DSO, Palmer M (2009) Online social capital: understanding e-impulse buying in practice. J Retail Consum Serv 16(4):320–328
24. Liu Y, Li H, Hu F (2013) Website attributes in urging online impulse purchase: an empirical investigation on consumer perceptions. Decis Support Syst 55(3):829–837
25. Madhavaram SR, Laverie DA (2004) Exploring impulse purchasing on the internet. Adv Consum Res 31:59–66

26. Martinez-Lopez FJ (2006) Motivations for consumption behaviours on the web: a conceptual model based on a holistic approach. Int J Electron Mark Retail 1(1):3–20
27. Mattila AS, Wirtz J (2008) The role of store environmental stimulation and social factors on impulse purchasing. J Serv Mark 22(7):562–567
28. Novak TP, Hoffman DL, Duhachek A (2003) The influence of goal-directed and experiential activities on online flow experiences. J Consum Psychol 13(1):3–16
29. Odekerken-Schröder G, Wulf KD, Schumacher P (2003) Strengthening outcomes of retailer-consumer relationships: the dual impact of relationship marketing tactics and consumer personality. J Bus Res 56(3):177–190
30. Park EJ, Kim EY, Forney JC (2006) A structural model of fashion-oriented impulse buying behavior. J Fashion Mark Manage 10(4):433–446
31. Piron F (1991) Defining impulse purchasing. Adv Consum Res 18(1):509–514
32. Sajad R, Faizan A et al (2015) Online impulse buying of tourism products: the role of web site personality, utilitarian and hedonic web browsing. J Hospitality Tourism Technol 7(1):60–83
33. Schiffman K (2005) Behavior del consumer, 8th edn. Prentice Hall, México
34. Sharma P, Sivakumaran B, Marshall R (2010) Impulse buying and variety seeking: a trait-correlates perspective. J Bus Res 63(3):276–283
35. Stern H (1962) The significance of impulse buying today. J Mark 26(2):59–62
36. Verhagen T, Dolen WV (2011) The influence of online store beliefs on consumer online impulse buying: a model and empirical application. Inf Manage 48(8):320–327
37. Wells JD, Parboteeah DV, Valacich JS (2011) Online impulse buying: understanding the interplay between consumer impulsiveness and website quality. J Assoc Inf Syst 12(1):32–56
38. Wu K, Vassileva J et al (2016) Complexity or simplicity designing product pictures for advertising in online marketplaces. J Retail Consum Serv 28:17–27

Decision Support System

Fuzzy Multi-attribute Grey Relational Analysis Using DEA and AHP

Mohammad Sadegh Pakkar(✉)

Faculty of Management, Laurentian University, Sudbury, ON P3E 2C6, Canada
ms_pakkar@laurentian.ca

Abstract. This research applies the method of grey relational analysis (GRA) for multiple attribute decision making (MADM) problems in which the attribute weights are completely unknown and attribute values take the form of fuzzy numbers. In order to obtain the attribute weights, this research proposes an integrated data envelopment analysis (DEA) and analytic hierarchy process (AHP) approach. According to this, we define two sets of weights in a domain of grey relational loss, i.e., a reduction in grey relational grade, between each alternative and the ideal alternative. The first set represents the weights of attributes with the minimal grey relational loss in DEA. The second set represents the priority weights of attributes, bounded by AHP, with the maximal grey relational loss. Using a parametric goal programming model, we explore the various sets of weights in a defined domain of grey relational loss. This may result in various ranking positions for each alternative in comparison to the other alternatives. An illustrated example of a nuclear waste dump site selection is used to highlight the usefulness of the proposed approach.

Keywords: Grey relational analysis · Data envelopment analysis · Analytic hierarchy process · Multiple attribute decision making · Goal programming · Fuzzy numbers

1 Introduction

Multiple attribute decision making (MADM) aims to find the ranking position of alternatives in the presence of multiple incommensurate attributes. Many MADM problems take place in an environment in which the information about attribute weights are incompletely known and attribute values take the form of intervals and fuzzy numbers [16,24,29].

Grey relational analysis (GRA) is part of grey system theory [3], which is suitable for solving a variety of MADM problems with both crisp and fuzzy data. The application of GRA with fuzzy data has recently attracted the attention of many scholars [5,8,25].

GRA solves MADM problems by aggregating incommensurate attributes for each alternative into a single composite value while the weight of each attribute is

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_57

subject to the decision maker's judgment. When such information is unavailable equal weights seem to be a norm. However, this is often the source of controversies for the final ranking results. Therefore, how to properly select the attribute weights is a main source of difficulty in the application of this technique. Fortunately, the development of modern operational research has provided us two excellent tools called data envelopment analysis (DEA) and analytic hierarchy process (AHP), which can be used to derive attribute weights in GRA.

DEA is an objective data-oriented approach to assess the relative performance of a group of decision making units (DMUs) with multiple inputs and outputs [2]. Traditional DEA models require crisp input and output data. However, in recent years, fuzzy set theory has been proposed as a way to quantify imprecise and vague data in DEA models [7,14,26]. In the field of GRA, DEA models without explicit inputs are applied, i.e., the models in which only pure outputs or index data are taken into account [11,27,28]. In these models, each DMU or alternative can freely choose its own favorable system of weights to maximize its performance. However, this freedom of choosing weights is equivalent to keeping the preferences of a decision maker out of the decision process.

Alternatively, AHP is a subjective data-oriented procedure which can reflect the relative importance of a set of attributes and alternatives based on the formal expression of the decision maker's preferences. AHP usually involves three basic functions: structuring complexities, measuring on a ratio-scale and synthesizing [23]. Some researchers incorporate fuzzy set theory in the conventional AHP to express the uncertain comparison judgments as fuzzy numbers [9,12,17].

However, AHP has been a target of criticism because of the subjective nature of the ranking process [4]. The application of AHP with GRA can be seen in [1,10,30].

In order to overcome the problematic issue of confronting the contradiction between the objective weights in DEA and subjective weights in AHP, this research proposes an integrated DEA and AHP approach in deriving the attribute weights in a fuzzy GRA methodology. This can be implemented by incorporating weight bounds using AHP in DEA-based GRA models. It is worth pointing out that the models proposed in this article are not brand-new models in the DEA-AHP literature. Conceptually, they are parallel to the application of DEA and AHP in GRA using crisp data as discussed in [20]. Nevertheless, it is the first time that these models are applied to a fuzzy GRA methodology. Further research on the integration of DEA and AHP approach in deriving the attribute weights with fuzzy data can be seen in [19].

2 Methodology

2.1 Fuzzy Multiple Attribute Grey Relational Analysis

Let $A = \{A_1, A_2, \cdots, A_m\}$ be a discrete set of alternatives and $C = \{C_1, C_2, \cdots, C_n\}$ be a set of attributes. Let $\tilde{y}_{ij} = (y_{1ij}, y_{2ij}, y_{3ij}, y_{4ij})$ be a trapezoidal fuzzy number representing the value of attribute $C_j (j = 1, 2, \cdots, n)$ for alternative $A_i (i = 1, 2, \cdots, m)$. Using α-cut technique, a trapezoidal fuzzy number can be transformed into an interval number as follows:

$$y_{ij} = [y_{ij}^-, y_{ij}^+] = [\alpha y_{2ij} + (1 - \alpha) y_{1ij}, \alpha y_{3ij} + (1 - \alpha) y_{4ij})], \tag{1}$$

where $y_{ij} = [y_{ij}^-, y_{ij}^+], y_{ij}^- \leq y_{ij}^+$, is an interval number representing the value of attribute $C_j (j = 1, 2, \cdots, n)$ for alternative $A_i (i = 1, 2, \cdots, m)$. Then alternative A_i is characterized by a vector $Y_i = \left([y_{i1}^-, y_{i1}^+], [y_{i2}^-, y_{i2}^+], \cdots, [y_{in}^-, y_{in}^+]\right)$ of attribute values. The term Y_i can be translated into the comparability sequence $R_i = ([r_{i1}^-, r_{i1}^+], [r_{i2}^-, r_{i2}^+], \cdots, [r_{in}^-, r_{in}^+])$ by using the following equations [31]:

$$[r_{ij}^-, r_{ij}^+] = \left[\frac{y_{ij}^-}{y_{j(\max)}^+}, \frac{y_{ij}^+}{y_{j(\max)}^+}\right] \quad \forall j, \quad y_{j(\max)}^+ = \max\{y_{1j}^+, y_{2j}^+, \cdots, y_{mj}^+\} \tag{2}$$

for desirable attributes,

$$[r_{ij}^-, r_{ij}^+] = \left[\frac{y_{i(\min)}^-}{y_{ij}^-}, \frac{y_{i(\min)}^-}{y_{ij}^+}\right] \quad \forall j, \quad y_{j(\min)}^- = \min\{y_{1j}^-, y_{2j}^-, \cdots, y_{mj}^-\} \tag{3}$$

for undesirable attributes.

Now, let A_0 be a virtual ideal alternative which is characterized by a reference sequence $U_0 = \left([u_{01}^-, u_{01}^+], [u_{02}^-, u_{02}^+], \cdots, [u_{0n}^-, u_{0n}^+]\right)$ of the maximum attribute values as follows:

$$u_{0j}^- = \max\{r_{1j}^-, r_{2j}^-, \cdots, r_{mj}^-\} \ \ \forall j, \tag{4}$$
$$u_{0j}^+ = \max\{r_{1j}^+, r_{2j}^+, \cdots, r_{mj}^+\} \ \ \forall j. \tag{5}$$

To measure the degree of similarity between $r_{ij} = [r_{ij}^-, r_{ij}^+]$ and $u_{0j} = [u_{0j}^-, u_{0j}^+]$ for each attribute, the grey relational coefficient, ξ_{ij}, can be calculated as follows:

$$\xi_{ij} = \frac{\min_i \min_j \left|[u_{0j}^-, u_{0j}^+] - [r_{ij}^-, r_{ij}^+]\right| + \rho \max_i \max_j \left|[u_{0j}^-, u_{0j}^+] - [r_{ij}^-, r_{ij}^+]\right|}{\left|[u_{0j}^-, u_{0j}^+] - [r_{ij}^-, r_{ij}^+]\right| + \rho \max_i \max_j \left|[u_{0j}^-, u_{0j}^+] - [r_{ij}^-, r_{ij}^+]\right|}, \tag{6}$$

while the distance between $u_{0j} = [u_{0j}^-, u_{0j}^+]$ and $r_{ij} = [r_{ij}^-, r_{ij}^+]$ is measured by $|u_{0j} - r_{ij}| = \max\left(\left|u_{0j}^- - r_{ij}^-\right|, \left|u_{0j}^+ - r_{ij}^+\right|\right)$. $\rho \in [0, 1]$ is the distinguishing coefficient, generally $\rho = 0.5$. It should be noted that the final results of GRA for MADM problems are very robust to changes in the values of ρ. Therefore, selecting the different values of ρ would only slightly change the rank order of attributes [13]. To find an aggregated measure of similarity between alternative A_i, characterized by the comparability sequence R_i, and the ideal alternative A_0, characterized by the reference sequence U_0, over all the attributes, the grey relational grade, Γ_i, can be computed as follows:

$$\Gamma_i = \sum_{j=1}^{n} w_j \xi_{ij}, \tag{7}$$

where w_j is the weight of attribute C_j and $\sum_{j=1}^{n} w_j = 1$. In practice, expert judgments are often used to obtain the weights of attributes. When such information is unavailable equal weights seem to be a norm. Nonetheless, the use of equal weights does not place an alternative in the best ranking position in comparison to the other alternatives. In the next section, we show how DEA can be used to obtain the optimal weights of attributes for each alternative in GRA.

2.2 DEA-Based GRA Models

Since all the grey relational coefficients are benefit (output) data, a DEA-based GRA model can be formulated similar to a classical DEA model without explicit inputs [15]:

$$\Gamma_k = \max \sum_{j=1}^{n} w_j \xi_{kj}, \tag{8}$$

$$\sum_{j=1}^{n} w_j \xi_{ij} \leq 1 \quad \forall i, \tag{9}$$

$$w_j > 0 \quad \forall j, \tag{10}$$

where Γ_k is the grey relational grade for alternative under assessment A_k (known as a decision making unit in the DEA terminology). k is the index for the alternative under assessment where k ranges over $1, 2, \cdots, m$. w_j is the weight of attribute C_j. The first set of constraints (9) assures that if the computed weights are applied to a group of m alternatives, $(i = 1, 2, \cdots, m)$, they do not attain a grade of larger than 1. The process of solving the model is repeated to obtain the optimal grey relational grade and the optimal weights required to attain such a grade for each alternative. The objective function (8) in this model maximizes the ratio of the grey relational grade of alternative A_k to the maximum grey relational grade across all alternatives for the same set of weights $(\max \Gamma_k / \max_{i=1,\cdots,m} \Gamma_i)$. Hence, an optimal set of weights in the DEA based-GRA model represents A_k in the best light in comparison to all the other alternatives. It should be noted that the grey relational coefficients are normalized data. Consequently, the weights attached to them are also normalized. In addition, adding the constraint $\sum_{j=1}^{n} w_j = 1$ to the DEA-based GRA model is not recommended here. In fact, the sum-to-one constraint is a non-homogeneous constraint (i.e., its right-hand side is a non-zero free constant) which can lead to underestimation of the grey relational grades of alternatives or infeasibility in the DEA-based GRA model (see [22])

2.3 Minimax DEA-Based GRA Model Using AHP

We develop our formulation based on a simplified version of the generalized distance model (see for example [6]). Let $\Gamma_k^* (k = 1, 2, \cdots, m)$ be the best attainable grey relational grade for the alternative under assessment, calculated from the

DEA-based GRA model. We want the grey relational grade, $\Gamma_k(w)$, calculated from the vector of weights $w = (w_1, \cdots, w_n)$ to be closest to Γ_k^*. Our definition of "*closest*" is that the largest distance is at its minimum. Hence we choose the form of the minimax model: $\min_w \max_k \{\Gamma_k^* - \Gamma_k(w)\}$ to minimize a single deviation which is equivalent to the following linear model:

$$\min \ \theta \tag{11}$$

$$\text{s. t. } \Gamma_k^* - \sum_{j=1}^{n} w_j \xi_{kj} \le \theta, \tag{12}$$

$$\sum_{j=1}^{n} w_j \xi_{ij} \le \Gamma_i^* \quad \forall i, \tag{13}$$

$$\theta \le 1, \tag{14}$$

$$\theta, w_j \ge 0 \quad \forall j. \tag{15}$$

The combination of (11)–(15) forms a minimax DEA based-GRA model that identifies the minimum grey relational loss $\theta_{\min}$ needed to arrive at an optimal set of weights. The first constraint ensures that each alternative loses no more than θ of its best attainable relational grade, Γ_k^*. The second set of constraints satisfies that the relational grades of all alternatives are less than or equal to their upper bound of Γ_k^*. It should be noted that for each alternative, the minimum grey relational loss $\theta = 0$. Therefore, the optimal set of weights obtained from the minimax DEA based-GRA model is exactly similar to that obtained from the DEA-based GRA model.

On the other hand, the priority weights of attributes are defined out of the internal mechanism of DEA by AHP. In order to more clearly demonstrate how AHP is integrated into the newly proposed minimax DEA-based GRA model, this research presents an analytical process in which attributes' weights are bounded by the AHP method. The AHP procedure for imposing weight bounds may be broken down into the following steps:

Step 1. A decision maker makes a pairwise comparison matrix of different attributes, denoted by B with the entries of $b_{hq} (h = q = 1, 2, \cdots, n)$. The comparative importance of attributes is provided by the decision maker using a rating scale. Saaty [23] recommends using a 1–9 scale.

Step 2. The AHP method obtains the priority weights of attributes by computing the eigenvector of matrix B (Eq. (16)), $e = (e_1, e_2, \cdots, e_j)^T$, which is related to the largest eigenvalue, $\lambda_{\max}$.

$$Be = \lambda_{\max} e. \tag{16}$$

To determine whether or not the inconsistency in a comparison matrix is reasonable the random consistency ratio, $C.R.$ can be computed by the following equation:

$$C.R. = \frac{\lambda_{\max} - N}{(N-1)R.I.}, \tag{17}$$

where $R.I.$ is the average random consistency index and N is the size of a comparison matrix.

In order to estimate the maximum relational loss $\theta_{\max}$ necessary to achieve the priority weights of attributes for each alternative, the following set of constraints is added to the minimax DEA-based GRA model:

$$w_j = \alpha e_j \quad \forall j. \tag{18}$$

The set of constraints (18) changes the priority weights of attributes to weights for the new system by means of a scaling factor α. The scaling factor α is added to avoid the possibility of contradicting constraints leading to infeasibility or underestimating the grey relational grade of alternatives (see [22]).

2.4 A Parametric Goal Programming Model

In this stage we develop a parametric goal programming model that can be solved repeatedly to generate the various sets of weights for the discrete values of the parameter θ, such that $0 \le \theta \le \theta_{\max}$. Let $w(\theta)$ be a vector of attribute weights for a given value of parameter θ. Let $w^*(\theta_{\max})$ be the vector of priority weights of attributes obtained from the minimax DEA-based GRA model after adding the set of constraints (18). Our objective is to minimize the total deviations between $w(\theta)$ and $w^*(\theta_{\max})$ with a city block distance measure. Choosing such a distance measure, each deviation is being equally weighted subject to the following constraints:

$$\min \ Z_k(\theta) = \sum_{j=1}^{n} (d_j^+ + d_j^-) \tag{19}$$

$$\text{s. t. } w_j - d_j^+ + d_j^- = \alpha e_j \quad \forall j, \tag{20}$$

$$d_j^+, d_j^- \ge 0. \tag{21}$$

and constraints (12)–(15), where d_j^+ and d_j^- are the positive and negative deviations from the priority weight of attribute $C_j (j = 1, 2, \cdots, n)$ for alternative $A_k (k = 1, 2, \cdots, m)$. The set of Eq. (20) indicates the goal equations whose right-hand sides are the priority weights of attributes adjusted by a scaling variable.

Because the range of deviations computed by the objective function is different for each alternative, it is necessary to normalize it by using relative deviations rather than absolute ones. Hence, the normalized deviations can be computed by:

$$\Delta_k(\theta) = \frac{Z_k^*(0) - Z_k^*(\theta)}{Z_k^*(0)}, \tag{22}$$

where $Z_k^*(\theta)$ is the optimal value of the objective function for $0 \le \theta \le \theta_{\max}$. We define $\Delta_k(\theta)$ as a *measure of closeness* which represents the relative closeness of each alternative to the weights obtained from the minimax DEA-based GRA model in the range [0, 1] after adding the set of constraint (18) to it. Increasing the parameter (θ), we improve the deviations between the two systems of weights obtained from the minimax DEA-based GRA model before and after adding the set of constraints (18). This may lead to different ranking positions for each alternative in comparison to the other alternatives. It should be noted that in a special case where the parameter $\theta = \theta_{\max} = 0$, we assume $\Delta_k(\theta) = 1$.

3 Numerical Example: Nuclear Waste Dump Site Selection

In this section we present the application of the proposed approach for nuclear waste dump site selection. The multiple attribute data, adopted from Wu and Olson [27], are presented in Table 1. There are twelve alternative sites and 4 performance attributes. *Cost*, *Lives lost*, and *Risk* are undesirable attributes and *Civic improvement* is a desirable attribute. *Cost* is in billions of dollars. *Lives lost* reflects expected lives lost from all exposures. *Risk* shows the risk of catastrophe (earthquake, flood, etc.) and *Civic improvement* is the improvement of the local community due to the construction and operation of each site. *Cost* and *Lives lost* are crisp values as outlined in Table 1, but *Risk* and *Civic improvement* have fuzzy data for each nuclear dump site.

We use the processed data as reported in [27]. First the trapezoidal fuzzy data are used to express linguistic data in Table 1. Using the α-cut technique, the raw data are expressed in fuzzy intervals as shown in Table 2. These data are turned into the comparability sequence by using the Eqs. (2) and (3). Each attribute is now on a common 0–1 scale where 0 represents the worst imaginable attainment on an attribute, and 1.00 the best possible attainment.

Table 3 shows the results of a pairwise comparison matrix in the AHP model as constructed by the author in Expert Choice software. The priority weight for each attribute would be the average of the elements in the corresponding row of the normalized matrix of pairwise comparison, shown in the last column of Table 3. One can argue that the priority weights of attributes must be judged by nuclear safety experts. However, since the aim of this section is just to show the application of the proposed approach on numerical data, we see no problem to use our judgment alone.

Using Eq. (6), all grey relational coefficients are computed to provide the required (output) data for the DEA-based GRA model as shown in Table 4. Note that grey relational coefficients depend on the distinguishing coefficient ρ, which here is 0.80.

Solving the minimax DEA-based GRA model for the site under assessment, we obtain an optimal set of weights with minimum grey relational loss ($\theta_{\min}$). It should be noted that the value of the grey relational grade of all waste dump sites calculated from the minimax DEA-based GRA model is identical to that calculated from the DEA-based GRA model. Therefore, the minimum grey relational loss for the site under assessment is $\theta_{\min} = 0$ (Table 5). This implies that

Table 1. Data for nuclear waste dump site selection

Site	Cost	Lives	Risk	Civic
Nome	40	60	Very high	Low
Newark	100	140	Very low	Very high
Rock Sprgs	60	40	Low	High
Duquesne	60	40	Medium	Medium
Gary	70	80	Low	Very high
Yakima	70	80	High	Medium
Turkey	60	70	High	High
Wells	50	30	Medium	Medium
Anaheim	90	130	Very high	Very low
Epcot	80	120	Very low	Very low
Duckwater	80	70	Medium	Low
Santa Cruz	90	100	Very high	Very low

Table 2. Fuzzy interval nuclear waste dump site data

Site	Cost	Lives lost	Risk	Civic
Nome	[0.80–1.00]	[0.40–0.70]	[0.00–0.10]	[0.10–0.30]
Newark	[0.00–0.05]	[0.00–0.05]	[0.90–1.00]	[0.90–1.00]
Rock Sprgs	[0.70–0.95]	[0.70–0.90]	[0.70–0.90]	[0.70–0.90]
Duquesne	[0.50–0.85]	[0.70–0.90]	[0.40–0.60]	[0.40–0.60]
Gary	[0.40–0.60]	[0.10–0.30]	[0.70–0.90]	[0.90–1.00]
Yakima	[0.50–0.70]	[0.10–0.30]	[0.10–0.30]	[0.40–0.60]
Turkey	[0.75–0.90]	[0.20–0.40]	[0.10–0.30]	[0.70–0.90]
Wells	[0.85–0.95]	[0.85–1.00]	[0.40–0.60]	[0.40–0.60]
Anaheim	[0.00–0.30]	[0.00–0.10]	[0.00–0.10]	[0.00–0.10]
Epcot	[0.10–0.40]	[0.00–0.20]	[0.90–1.00]	[0.00–0.10]
Duckwater	[0.30–0.50]	[0.20–0.40]	[0.40–0.60]	[0.10–0.30]
Santa Cruz	[0.10–0.40]	[0.10–0.30]	[0.00–0.10]	[0.00–0.10]

Table 3. Pairwise comparison matrix of 4 attributes

Attribute	Cost	Lives	Risk	Civic	Priority
Cost	1	1/5	1/2	3	0.131
Lives	5	1	2	9	0.545
Risk	2	1/2	1	6	0.275
Civic	1/3	1/9	1/6	1	0.05
$C.R. = 0.01$					

Table 4. Results of grey relational coefficient for nuclear waste dump site selection

Site	Cost	Lives lost	Risk	Civic
Nome	0.9383	0.6281	0.4578	0.4872
Newark	0.4444	0.4444	1	1
Rock Sprgs	0.8352	0.8352	0.7917	0.7917
Duquesne	0.6847	0.8352	0.6032	0.6032
Gary	0.6281	0.5033	0.7917	1
Yakima	0.6847	0.5033	0.4872	0.6032
Turkey	0.8837	0.539	0.4872	0.7917
Wells	0.9383	1	0.6032	0.6032
Anaheim	0.472	0.4578	0.4578	0.4578
Epcot	0.5033	0.472	1	0.4578
Duckwater	0.5802	0.539	0.6032	0.4872
Santa Cruz	0.5033	0.5033	0.4578	0.4578

the measure of relative closeness to the AHP weights for the site under assessment is $\varDelta_k(\theta_{\min}) = 0$. On the other hand, solving the minimax DEA-based GRA model for the site under assessment after adding the set of constraints (18), we adjust the priority weights of attributes (outputs) obtained from AHP in such a way that they become compatible with the weights' structure in the minimax DEA-based GRA model. This results in the maximum grey relational loss, $\theta_{\max}$, for the site under assessment (Table 5). In addition, this implies that the measure of relative closeness to the AHP weights for the site under assessment is $\varDelta_k(\theta_{\max}) = 1$.

Table 6 presents the optimal weights of attributes as well as its scaling factor for all nuclear waste dump sites. It should be noted that the priority weights of AHP (Table 3) used for incorporating weight bounds on the attribute weights are obtained as $e_j = \frac{w_j}{\alpha}$.

Going one step further to the solution process of the parametric goal programming model, we proceed to the estimation of total deviations from the AHP weights for each site while the parameter θ is $0 \leq \theta \leq \theta_{\max}$. Table 7 represents the ranking position of each site based on the minimum deviation from the priority weights of attributes for $\theta = 0$. It should be noted that in a special case where the parameter $\theta = \theta_{\max} = 0$, we assume $\varDelta_k(\theta) = 0$. Table 7 shows that Wells is the best alternative in terms of the grey relational grade and its relative closeness to the priority weights of attributes.

Nevertheless, increasing the value of θ from 0 to $\theta_{\max}$ has two main effects on the performance of the other sites: improving the degree of deviations and reducing the value of the grey relational grade. This, of course, is a phenomenon, one expects to observe frequently. The graph of $\varDelta(\theta)$ versus θ, as shown in Fig. 1, is used to describe the relation between the relative closeness to the priority weights of attributes, versus the grey relational loss for each site. This

Table 5. Minimum and maximum grey relational losses for each nuclear waste dump site

Site	Γ_k^*	$\theta_{\min}$	$\theta_{\max}$
Nome	1	0	0.2876
Newark	1	0	0.2761
Rock Sprgs	1	0	0.0487
Duquesne	0.8921	0	0.0345
Gary	1	0	0.2774
Yakima	0.7855	0	0.1742
Turkey	1	0	0.3251
Wells	1	0	0.0000
Anaheim	0.5735	0	0.0409
Epcot	1	0	0.2811
Duckwater	0.7351	0	0.0869
Santa Cruz	0.5943	0	0.0283

Table 6. Optimal weights of minimax DEA-based GRA model for all nuclear waste dump sites bounded by AHP

w_1	w_2	w_3	w_4	α
0.1516	0.6308	0.3183	0.0579	1.1575

Table 7. The ranking position of each site based on the minimum distance to priority weights of attributes

Site	$Z^*(\theta_{\min} = 0)$	Rank	Site	$Z^*(\theta_{\min} = 0)$	Rank
Nome	1.1568	12	Turkey	0.8884	9
Newark	0.3955	4	Wells	0.0000	1
Rock Sprgs	0.1479	2	Anaheim	0.7521	7
Duquesne	0.2069	3	Epcot	0.5274	5
Gary	0.7049	6	Duckwater	0.8371	8
Yakima	0.9496	11	Santa Cruz	0.8972	10

may result in different ranking positions for each site in comparison to the other sites. In order to clearly discover the effect of grey relational loss on the ranking position of each nuclear dump site, as shown in Table 8 in Appendix, we performed a Kruskal-Wallis test. The Kruskal-Wallis test compares the medians of rankings to determine whether there is a significant difference between them. The result of the test reveals that its p-value is quite smaller than 0.01. Therefore, we conclude that increasing grey relational loss in the whole range [0.0, 0.33] changes the ranking position of each site significantly. Note that at

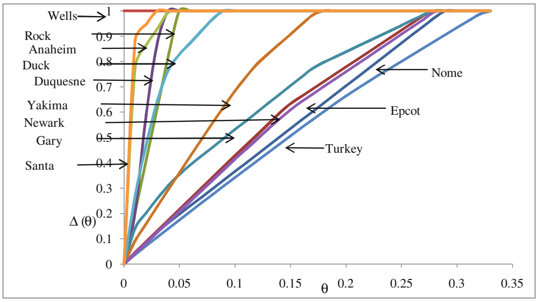


Fig. 1. The relative closeness to the priority weights of attributes [$\Delta(\theta)$], versus grey relational loss (θ) for each site

$\theta = 0$ sites can be ranked based on $Z_k^*(0)$ from the closest to the furthest from the priority weights of attributes. For instance, at $\theta = 0$, Nome, Newark and Rock Sprgs with grey relational grades of one, are ranked in 12^{th}, 4^{th} and 2^{nd} places, respectively (Tables 5 and 7). However, with a small grey relational loss at $\theta = 0.01$, Nome, Newark and Rock Sprgs take 9^{th}, 10^{th} and 5^{th} places in the rankings, respectively. Using this example, as a guideline, it is relatively easy to rank the sites in terms of distance to the priority weights of attributes. At $\theta = 0.02$, Newark moves up into 9^{th} place while Nome and Rock Sprgs drop in 10^{th} and 6^{th} places, respectively. It is clear that both measures, $Z_k^*(0)$ and $\Delta_k(\theta)$, are necessary to explain the ranking position of each nuclear dump site.

4 Conclusion

We develop an integrated approach based on DEA and AHP methodologies for deriving the attribute weights in GRA with fuzzy data. We define two sets of attribute weights in a minimax DEA-based GRA framework. The first set represents the weights of attributes with minimum grey relational loss. The second set represents the corresponding priority weights of attributes, using AHP, with maximum grey relational loss. We assess the performance of each alternative (or DMU) in comparison to the other alternatives based on the relative closeness of the first set of weights to the second set of weights. Improving the measure of relative closeness in a defined range of grey relational loss, we explore the various ranking positions for the alternative under assessment in comparison to the other alternatives. To demonstrate the effectiveness of the proposed approach, an illustrative example of a nuclear waste dump site using twelve alternative sites and 4 attributes is carried out. Further studies can apply the simultaneous application of DEA and AHP to the field of GRA by considering the hierarchical structures of attributes in the ranking positions of alternatives [18, 21].

Appendix

Table 8. The measure of relative closeness to the priority weights of attributes $[\Delta_k(\theta)]$ verses grey relational loss$[\theta]$ for each nuclear waste dump site

θ	Nome	Newark	Rock Sprgs	Duquesne	Gary	Yakima	Turkey	Wells	Anaheim	Epcot	Duckwater	Santa Cruz
0	0	0	0	0	0	0	0	0	0	0	0	0
Rank	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
0.01	0.0451	0.0445	0.2054	0.2051	0.1409	0.0921	0.0369	1.0000	0.7774	0.0418	0.2693	0.8839
Rank	9	10	5	6	7	8	12	1	3	11	4	2
0.02	0.0870	0.0882	0.4108	0.5901	0.2019	0.1596	0.0725	1.0000	0.8644	0.0835	0.4711	0.9474
Rank	10	9	6	4	7	8	12	1	3	11	5	2
0.03	0.1263	0.1310	0.6162	0.8726	0.2586	0.2270	0.1068	1.0000	0.9292	0.1251	0.6439	1.0000
Rank	10	9	6	4	7	8	12	1	3	11	5	2
0.04	0.1627	0.1734	0.8217	1.0000	0.3100	0.2941	0.1405	1.0000	0.9940	0.1667	0.7656	1.0000
Rank	11	9	5	3	7	8	12	1	4	10	6	1
0.05	0.1967	0.2158	1.0000	1.0000	0.3550	0.3610	0.1743	1.0000	1.0000	0.2082	0.8284	1.0000
Rank	11	9	5	1	8	7	12	1	4	10	6	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.29	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9070	1.0000	1.0000	1.0000	1.0000	1.0000
Rank	1	1	1	1	1	1	12	1	1	1	1	1
0.3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9335	1.0000	1.0000	1.0000	1.0000	1.0000
Rank	1	1	1	1	1	1	12	1	1	1	1	1
0.31	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9600	1.0000	1.0000	1.0000	1.0000	1.0000
Rank	1	1	1	1	1	1	12	1	1	1	1	1
0.32	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9865	1.0000	1.0000	1.0000	1.0000	1.0000
Rank	1	1	1	1	1	1	12	1	1	1	1	1
0.33	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Rank	1	1	1	1	1	1	1	1	1	1	1	1

References

1. Birgün S, Güngör C (2014) A multi-criteria call center site selection by hierarchy grey relational analysis. J Aeronaut Space Technol 7(1):1–8
2. Cooper WW, Seiford LM, Zhu J (2011) Handbook on data envelopment analysis. Springer, US
3. Deng JL (1982) Control problems of grey systems. Syst Control Lett 1(5):288–294
4. Dyer JS (1990) Remarks on the analytic hierarchy process. Manage Sci 36(3):249–258
5. Goyal S, Grover S (2012) Applying fuzzy grey relational analysis for ranking the advanced manufacturing systems. Grey Syst 2(2):284–298
6. Hashimoto A, Wu DA (2004) A DEA-compromise programming model for comprehensive ranking. J Oper Res Soc Jpn 2(2):73–81
7. Hatami-Marbini A, Saati S, Tavana M (2010) Data envelopment analysis with fuzzy parameters: an interactive approach. Core Discuss Pap Rp 2(3):39–53
8. Hou J (2010) Grey relational analysis method for multiple attribute decision making in intuitionistic fuzzy setting. J Converg Inf Technol 5(10):194–199
9. Javanbarg MB, Scawthorn C et al (2012) Fuzzy AHP-based multicriteria decision making systems using particle swarm optimization. Expert Syst Appl 39(1):960–966

10. Jia W, Li C, Wu X (2011) Application of multi-hierarchy grey relational analysis to evaluating natural gas pipeline operation schemes. Springer, Heidelberg
11. Jun G, Xiaofei C (2013) A coordination research on urban ecosystem in Beijing with weighted grey correlation analysis based on DEA. J Appl Sci 13(24):5749–5752
12. Kahraman C, Cebeci U, Ulukan Z (2003) Multi-criteria supplier selection using fuzzy AHP. Logist Inf Manage 16(6):382–394
13. Kuo Y, Yang T, Huang GW (2008) The use of grey relational analysis in solving multiple attribute decision-making problems. Comput Ind Eng 55(1):80–93
14. Lertworasirikul S, Fang SC et al (2003) Fuzzy data envelopment analysis (DEA): a possibility approach. Fuzzy Sets Syst 139(2):379–394
15. Liu WB, Zhang DQ et al (2011) A study of DEA models without explicit inputs. Omega 39(5):472–480
16. Olson DL, Wu D (2006) Simulation of fuzzy multiattribute models for grey relationships. Eur J Oper Res 175(1):111–120
17. Osman MS, Gadalla MH et al (2013) Fuzzy analytic hierarchical process to determine the relative weights in multi-level programming problems. Int J Math Arch 4(7):282–295
18. Pakkar MS (2015) An integrated approach based on DEA and AHP. CMS 12(1):153–169
19. Pakkar MS (2016) An integrated approach to grey relational analysis, analytic hierarchy process and data envelopment analysis. J Cent Cathedra 9(1):71–86
20. Pakkar MS (2016) Multiple attribute grey relational analysis using DEA and AHP. Complex Intell Syst 2(4):243–250
21. Pakkar MS (2016) Using DEA and AHP for hierarchical structures of data. Ind Eng Manage Syst 15(1):49–62
22. Podinovski VV (2004) Suitability and redundancy of non-homogeneous weight restrictions for measuring the relative efficiency in DEA. Eur J Oper Res 154(2):380–395
23. Saaty RW (1987) The analytic hierarchy process - what it is and how it is used. Math Modell 9(3–5):161–176
24. Wei G, Wang H et al (2011) Grey relational analysis method for intuitionistic fuzzy multiple attribute decision making with preference information on alternatives. Int J Comput Intell Syst 4(2):164–173
25. Wei GW (2010) GRA method for multiple attribute decision making with incomplete weight information in intuitionistic fuzzy setting. Knowl-Based Syst 23(3):243–247
26. Wen M, Li H (2009) Fuzzy data envelopment analysis (DEA): model and ranking method. J Comput Appl Math 223(2):872–878
27. Wu DD, Olson DL (2010) Fuzzy multiattribute grey related analysis using DEA. Comput Math Appl 60(1):166–174
28. Xie Z, Mo L (2013) Analysis method and its application of weighted grey relevance based on super efficient DEA. Res J Appl Sci Eng Technol 5(2):470–474
29. Xu Z (2005) On method for uncertain multiple attribute decision making problems with uncertain multiplicative preference information on alternatives. Fuzzy Optim Decis Making 4(2):131–139
30. Zeng G, Jiang R et al (2007) Optimization of wastewater treatment alternative selection by hierarchy grey relational analysis. J Environ Manage 82(2):250–9
31. Zhang J, Wu D, Olson DL (2005) The method of grey related analysis to multiple attribute decision making problems with interval numbers. Math Comput Modell 42(9):991–998

A Fuzzy Multi-criteria Decision Making Approach for Supplier Selection Under Fuzzy Environment

Adnan Sarwar[1], Ziqiang Zeng[1,2](✉), Richard AduAgyapong[1], Nguyen ThiHoaiThuong[1,3], and Talat Qadeer[1]

[1] Uncertainty Decision-Making Laboratory, Sichuan University, Chengdu 610064, People's Republic of China
zengziqiang@scu.edu.cn
[2] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA
[3] Faculty of Financial Accounting and Business Management, Vietnam National University of Agriculture, Hanoi, Vietnam

Abstract. Supplier selection is crucial and multi-criteria decision making problem in supply chain. Nowadays in competitive business environment, supplier selection is a critical task for purchasing department in every organization. Appropriate supplier helps manufacturer to reduce cost, consistent quality product, and enhance competitiveness in market. In order to select potential supplier, it is essential to tradeoff between tangible and intangible factors. Uncertainty and vagueness of decision makers opinion is an important characteristic of this problem. Fuzzy analytical hierarchical process (AHP) and extent analysis method is used to choose appropriate supplier under fuzzy environment. The application procedure of FAHP and extent analysis method is elaborated through numerical example.

Keywords: Supplier selection · Multi-criteria decision making · Fuzzy set theory · Fuzzy AHP · Extent analysis

1 Introduction

Supplier selection is the most challenging issue in supply chain. In order to develop sustainable supply chain companies need to select right supplier. Appropriate selection of supplier helps companies to provide high quality product at right price and at right time. The term "Supply chain management" (SCM) can be defined as it is a process that includes: sourcing raw material, convert raw material to finish goods, and delivers product to final customer. The purpose of SCM is to reduce inventory and production cycle time, to increase production as well as achieve long terms goal of firms with respect to customer stratifications. Appropriate selection of supplier is important task of purchasing department in

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_58

any organization. In order to select supplier, it is necessary to evaluate supplier based on different criteria. The supplier selection problem is considered as a multi-criteria decision making problem. The process of supplier evaluation and selection consists both quantitative and qualitative criteria [3]. Selection of supplier against single factor is not sufficient. Supplier should be evaluated against multi-criterion [11]. In supplier selection problem, it is necessary to be tradeoff between tangible and intangible factors which may conflict each other. Optimal selection of supplier can enable companies to manufacture innovative quality product and gain competitive advantage in market. An effective supplier selection process plays important role to success of any organization. Whether right or wrong selection of supplier directly effect on price, quality, and delivery time. The purpose of supplier selection is to minimize purchase risk and build strong relationship between supplier and buyer in competitive scenario [17]. Selection of potential supplier is strategic decision that helps firms to achieve goal for a longer period of time. Nowadays companies are changing reactive buying to proactive buying to select wealthy supplier [12].

Supplier selection problem can be categorized into two types. Single sourcing and multiple sourcing. Single sourcing means one supplier can provide all required material to the buyer. In multiple sourcing, single supplier unable to satisfy all needs of manufacturer. Hence companies need to split their required material among various supplier [7]. Dickson [8] ranked 23 different criteria includes quality, delivery, warranties, performance history, price, supplier profile, financial position, and technical capability for the purpose of supplier evaluation and selection. The author mentioned that the most important criteria for supplier selection is quality. Weber, et al. [26] reviewed 74 articles from the period 1966 to 1991 about supplier selection criteria and techniques. The author classified supplier selection approaches into three different types; linear weighting methods, mathematical programming models, and statistical techniques. De Boer, et al. [4] defined various phases of supplier selection process. In the initial phase, define problem, formulation of criteria, qualification of supplier, and selection of potential supplier among suppliers.

The process of supplier selection can be simple if one criteria is used in the process of decision making. However, in case of more than one criteria, buyers need to consider a range of different criteria in decision making process. If manufactures are used multiple criteria in supplier selection process, manufacture needs to consider influence of each criteria in decision making process, whether weights are equally given to each criteria or according to type and importance of different criteria [28]. In supplier selection process, two issues are very significant. Formulation of criteria that affect performance of supplier and determine methodology that can be used to select the best supplier. Therefore, selection of appropriate supplier is a critical task for any organization that affects organization efficiency and profitability. Every supplier has a number of strengths and weakness that should be assessed by buyer before ranking them.

2 Literature Review

Recently, supply chain management and supplier selection process have gained great attention in the literature. Appropriate supplier selection is a multi-criteria problem but traditionally supplier selection based on only one criteria price. The supplier selection process is one of the most important task for every firm to establish effective supply chain. In competitive environment, it is critical task for firms to select right supplier because potential supplier helps organizations to produce high quality product at reasonable price. In past, various methodologies have been used to select supplier [1]. Proposed integrated methodology data envelopment analysis (DEA) and group analytical hierarchy process (GAHP) to evaluate and select most efficient supplier. Chatterjee, et al. [12] proposed case based reasoning and decision support system including multi-attribute analysis to measure supplier management performance. Usually, supplier evaluation consists price, quality, flexibility, but in this paper some environment factors were considered in supplier selection process using knowledge-based system. Humphreys et al. [18] presented two methodologies for supplier selection problem and compare relatives performance of organization. Vikor and outranking methods were used to select supplier and compare the relative performance of supplier. Multi-criteria decision making model was applied in a construction industry for supplier evaluation and selection [21]. Multi-criteria decision making enabled construction industry to build good relationships among its supplier, managers, and partners.

The balance and ranking method were used to select supplier based on multi-criteria; profitability of supplier, technological capability, conformance quality, and relationship closeness. The application of model consists three steps; construction of out-ranking matrix, determine relative frequency of each supplier, then triangularised out-ranking matrix to obtain implicit order of each supplier with the help of balancing method [24]. The activity based costing (ABC) technique was proposed for evaluation and selection of vendor. The ABC technique helped to calculate total cost caused by vendor in a company's production process, to judge performance of supplier [19]. The multi-criteria group decision model for supplier selection was applied. This model based on ELECTRE IV and VIP analysis approach. The proposed model consists two stages. In first stage, ELECTRE IV utilized to determine ranking of different criteria. In second stage, VIP analysis method was used to select alternatives [2]. Dimensional analysis technique proposed to measure performance of supplier and to get index called vendor performance index (VPI). In the process of supplier performance evaluation both qualitative and quantitative criterion were used. To overcome blindness of human and vagueness fuzzy approach applied [16]. Supplier selection is crucial task for every organization. The analytical network process (ANP) method was proposed to select the best supplier based on three criteria, business structure of supplier, manufacture capability of supplier, and quality system of supplier. These main three criteria further classified to 45 sub-criteria [10,27]. Neural network (NN) approach used to select potential supplier. The determinant factors quality, performance history, geographical and price were selected in the process of supplier selection.

In Telecommunications Company AHP model was proposed for selection of vendor. Vendor selection is a complex and multi-criteria decision making problem. In application process of AHP, two strategic factors (criteria) cost and quality selected for evaluation of vendor, further both criteria categorized into 26 sub-criteria [23]. Shyura & Hsu-Shih [22] developed integrated model for vendor selection. The hybrid model was included analytical network process (ANP), TOPSIS, and nominal group.

A fuzzy multi-objective integer programming method was developed for supplier selection problem which incorporate three objectives cost minimization, maximize on-time delivery, and quality maximization [14]. Proposed strategy-aligned fuzzy simple multi-attribute rating technique (SMART) to solve supplier selection problem in supply chain. Strategy-aligned fuzzy SMART technique involves three stages. In first stage, define objective and strategy, second step, develop supply chain strategy, third stage recognize criteria for supplier selection [6, 15]. In this study, fuzzy analytical hierarchy process method was proposed with the concept of benefits, opportunities, cost, and risks (BOCR). With the application of this approach supplier can be evaluated various aspects based on quantitative and qualitative criteria. Fuzzy set theory was applied to overcome human ambiguity in decision making process.

In this paper, fuzzy AHP method is applied to select best supplier. In fuzzy AHP method, linguistic variables and triangular fuzzy numbers are used as a pairwise comparison scale to determine priority of main criteria and sub-criteria. After that, extent analysis method is applied to calculate final weight priority of main decision criteria, sub-criteria and alternatives.

3 Fuzzy AHP

The analytical hierarchical process (AHP) introduced by [20]. AHP is very useful method to solve multi-criteria decision making problem. It offers hierarchical procedure to solve problem using both subjective and objective measures. The hierarchical model of supplier selection problem consists four stages; objective, criteria, sub-criteria, and last stage is decision of alternatives.

The advantage of AHP is simple and ease of use but the most important disadvantage of this approach is that it uses nine-point scale which is unable to handle uncertainty in the process of comparisons of decision variables. Tradition AHP is unable to solve uncertain decision making problem, to overcome this problem, linguistic variables and triangular fuzzy numbers are used to decide priority of different decision variables. The extent analysis method is applied to calculate priority of weight using triangular fuzzy numbers. The fuzzy analytical hierarchical process is an extension of AHP which is used to handle fuzziness in supplier selection process based on both quantitative and qualitative criteria.

Fuzzy set theory and fuzzy number

Fuzzy set theory proposed by [29]. Fuzzy set theory helps to overcome vagueness of human, imprecise data, and uncertainty in decision making. Fuzzy set theory is used to represent mathematically human vagueness and uncertainty in decision making problem. A fuzzy set is a class of elements with continuum grades of membership. The membership value of each object lies between 0 and 1. Triangular fuzzy number is a part of fuzzy number whose membership functions are defined by three numbers [9]. Where parameter m donates the most promising value, l and u indicates lower and upper bond of fuzzy event.

The membership function $\mu_{\tilde{A}}$ is defined as

$$\mu_{\tilde{A}} = \left\{ \begin{array}{ll} 0, & x < 1 \\ (x - l/m - l)\,, & l \leq x \leq m, \\ (u - x/u - m)\,, & m \leq x \leq u \\ 0, & x > u \end{array} \right\}. \tag{1}$$

3.1 Extent Analysis for Fuzzy AHP

The extent analysis method introduced by [25]. In this paper extent analysis method is applied for supplier selection problem. In traditional AHP, nine point scale was used to determine priority of one criteria over another. In fuzzy AHP, linguistics variables and triangular fuzzy numbers are used. The fuzzy AHP based on extent analysis is applied for supplier selection problem. Let $X = \{x_1, x_2, \cdots, x_n\}$ represent object set of alternatives and $P = \{p_1, p_2, p_3, \cdots, p_n\}$ represent goal set of supplier selection. Each object is taken and extent analysis for each goal, p_1 is performed respectively [5]. Then m extent analysis value of each object can be obtained, using following signs:

$$M_{gi}^1, M_{gi}^2, \cdots, M_{gi}^m, i = 1, 2, \cdots, n, \tag{2}$$

where M_{gi}^j $(j = 1, 2, 3, \cdots, m)$ are Triangular fuzzy numbers. The steps of extent analysis method proposed by [25];

Step 1. The value of synthetic extent analysis with respect to i^{th} object is defined as

$$S_i = \sum_{j=1}^{m} M_{gi}^j \otimes \left[\sum_{i=1}^{n} \sum_{j=1}^{m} M_{gi}^j \right]^{-1}. \tag{3}$$

To obtain $\sum_{j=i}^{m} M_{gi}^j$, perform the fuzzy addition operation of m extent analysis values for a particular matrix such that,

$$\sum_{j=i}^{m} M_{gi}^j = \left(\sum_{j=1}^{m} l_j, \sum_{j=1}^{m} m_j, \sum_{j=1}^{m} u_j \right), \tag{4}$$

and to obtain $\left[\sum_{i=1}^{n}\sum_{j=1}^{m} M_{gi}^{j}\right]^{-1}$, perform the fuzzy addition operation of $M_{gi}^{j}\,(j = 1, 2, \cdots, m)$ values such that,

$$\sum_{i=1}^{n}\sum_{j=1}^{m} M_{gi}^{j} = \left(\sum_{i=1}^{n} l_i, \sum_{i=1}^{n} m_i, \sum_{i=1}^{n} u_i\right), \tag{5}$$

and then the inverse of the vector Eq. (5) is calculated such that,

$$\left[\sum_{i=1}^{n}\sum_{j=1}^{m} M_{gi}^{j}\right]^{-1} = \left(\frac{1}{\sum_{i=1}^{n} u_i}, \frac{1}{\sum_{i=1}^{n} m_i}, \frac{1}{\sum_{i=1}^{n} l_i}\right). \tag{6}$$

Step 2. The degree of possibility of $M_2 = (l_2, m_2, u_2) \geq M_1 = (l_1, m_1, u_1)$ is defined as

$$V\left(M_2 \geq M_1\right) = \sup_{y \geq x} \left\lfloor \min\left(\mu_{M_1}(x), \mu_{M_2}(y)\right)\right\rfloor \tag{7}$$

and can be equivalently expressed as follows;

$$V\left(M_2 \geq M_1\right) = hgt\left(M_1 \cap M_2\right) = \mu_{M_2}(d) = \left\{\begin{array}{ll} 1, & \text{if } m_2 \geq m_1 \\ 0, & \text{if } l_1 \geq u_2 \\ \frac{l_1 - u_2}{(m_2 - u_2) - (m_1 - l_1)}, & \text{otherwise} \end{array}\right\}, \tag{8}$$

where d is the ordinate of the highest intersection point D between μ_{M_1} and μ_{M_2}. In Fig. 1, the intersection between M_1 and M_2 has been shown.
To compare M_1 and M_2, we need both the values of $V\left(M_1 \geq M_2\right)$ and $V\left(M_2 \geq M_1\right)$.

Step 3. The degree of possibility for a convex fuzzy number to be greater than k convex fuzzy numbers $M_i\,(1, 2, \cdots, k)$ can be defined by

$$\begin{aligned} V\left(M \geq M_1, M_2, \cdots, M_k\right) &= V\left[\left(M \geq M_1\right) \text{ and } \left(M \geq M_2\right) \text{ and } \cdots \text{ and } \left(M \geq M_k\right)\right] \\ &= \min \mathrm{V}\left(M \geq M_i\right), \end{aligned} \tag{9}$$

Assume that

$$d'\left(A_i\right) = \min V\left(S_i \geq S_k\right). \tag{10}$$

For $k = 1, 2, \cdots, n;\ k \neq 1$.

$$W' = \left(d'\left(A_1\right), d'\left(A_2\right), \cdots, d'\left(A_n\right)\right)^T, \tag{11}$$

where $A_i\,(i = 1, 2, \cdots, n)$ are n elements.

Step 4. Via normalization, the normalized weight vectors are

$$W = \left(d\left(A_1\right), d\left(A_2\right), \cdots, d\left(A_n\right)\right)^T, \tag{12}$$

where W is a non-fuzzy numbers [25]. It gives the weight priority one alternative to another.

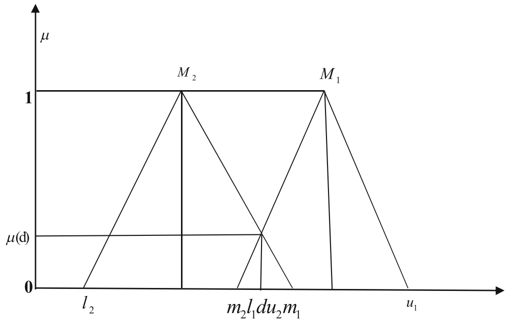


Fig. 1. The intersection between M2 and M1

4 Numerical Example

Suppose a manufacturing company wants to select supplier. Supplier selection is critical task for every company. The fuzzy AHP method is applied to select potential supplier. In supplier selection process, three suppliers have been selected for evaluation against three decision criteria. In evaluation process of supplier, each criteria is defined using linguistic judgments provided by decision makers (Table 1). The criteria is given below:

- Price: relates to product price and logistics management cost.
- Quality: relates to product durability and supporting service of quality.
- Flexibility: relates to production capacity flexibility and customization capability.

In Table 2, fuzzy pairwise comparison matrix of main decision criteria is made with respect to triangular fuzzy numbers. Fuzzy synthetic extent values Eq. (3) is used to calculate priority weight of main criteria (Fig. 2).

Table 1. Linguistic scale for supplier evaluation [13]

Linguistic variable	Fuzzy triangular number
Just equal	(1,1,1)
Equally preferred	(1/2,1,3/2)
Weakly more preferred	(1,3/2,2)
Strongly preferred	(3/2,2,5/2)
Very strongly preferred	(2,5/2,3)
Absolutely preferred	(5/2,3,7/2)

Table 2. Fuzzy pairwise comparison matrix of main attribute

	Price	Quality	Flexibility
Price	(1,1,1)	(1/2,1,3/2)	(1,3/2,2)
Quality	(2/3,1,2)	(1,1,1)	(3/2,2,5/2)
Flexibility	(1/2,2/3,1)	(2/5,1/2,2/3)	(1,1,1)

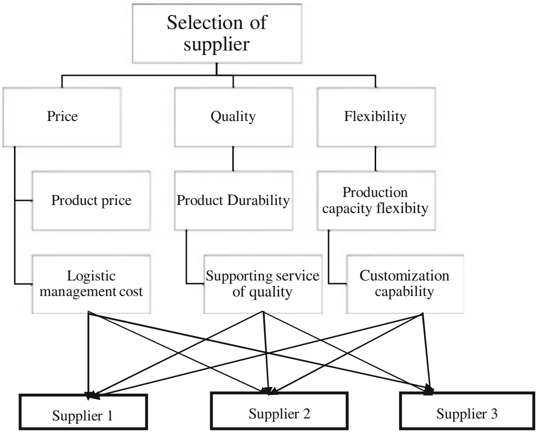


Fig. 2. Hierarchical structure of supplier selection problem

$$\text{Price} = (2.5, 3.5, 4.5) \otimes \left(\frac{1}{12.67}, \frac{1}{9.67}, \frac{1}{6.86}\right) = (0.20, 0.36, 0.66), \tag{13}$$

$$\text{Quality} = (3.17, 4, 5.5) \otimes \left(\frac{1}{12.67}, \frac{1}{9.67}, \frac{1}{6.86}\right) = (0.25, 0.41, 0.80), \tag{14}$$

$$\text{Flexibility} = (1.9, 2.17, 2.67) \otimes \left(\frac{1}{12.67}, \frac{1}{9.67}, \frac{1}{6.86}\right) = (0.15, 0.22, 0.39). \tag{15}$$

Next step is to calculate degree of possibility using Eqs. (7) and (8).

$$V(\text{Price} \geq \text{Quality}) = \frac{(0.25 - 0.66)}{(0.36 - 0.66) - (0.41 - 0.25)} = 0.89, \tag{16}$$

$$V(\text{Price} \geq \text{Flexibility}) = 1, \tag{17}$$

$$V(\text{Quality} \geq \text{Flexibility}) = 1, \tag{18}$$

$$V(\text{Quality} \geq \text{Price}) = 1, \tag{19}$$

$$V(\text{Flexibility} \geq \text{Price}) = \frac{(0.20 - 0.39)}{(0.22 - 0.39) - (0.36 - 0.20)} = 0.57, \tag{20}$$

$$V(\text{Flexibility} \geq \text{Quality}) = \frac{(0.25 - 0.39)}{(0.22 - 0.39) - (0.41 - 0.25)} = 0.42. \tag{21}$$

Table 3. Fuzzy pairwise comparison matrix with respect to price

	Product price	Logistics management cost
Product price	(1,1,1)	(3/2,2,5/2)
Logistics management cost	(2/5,1/2,2/3)	(1,1,1)

Normalized weight vector is $W_P = (1, 0)$.

Table 4. Fuzzy pairwise comparison matrix with respect to quality

	Product durability	Supporting service quality
Product durability	(1,1,1)	(1,3/2,2)
Supporting service quality	(1/2,2/3,1)	(1,1,1)

Normalized weight vector is $W_Q = (0.68, 0.32)$.

Table 5. Fuzzy pairwise comparison matrix with respect to flexibility

	Production capacity flexibility	Customization flexibility
Production capacity flexibility	(1,1,1)	(2,5/2,3)
Customization flexibility	(1/3,2/5,1/2)	(1,1,1)

Normalized weight vector is $W_F = (1, 0)$.

Table 6. Fuzzy pairwise comparison matrix of alternative with respect to product price

	Supplier 1	Supplier 2	Supplier 3
Supplier 1	(1,1,1)	(3/2,2,5/2)	(1,3/2,2)
Suupplier 2	(2/5,1/2,2/3)	(1,1,1)	(1/2,1,3/2)
Supplier 3	(1/2,2/3,1)	(2/3,1,2)	(1,1,1)

Normalized weight vector is $W_{PP} = (0.51, 0.20, 0.29)$.

Table 7. Fuzzy pairwise comparison matrix of alternative with respect to logistics management cost

	Supplier 1	Supplier 2	Supplier 3
Supplier 1	(1,1,1)	(2/3,1,2)	(1,3/2,2)
Supplier 2	(1/2,1,3/2)	(1,1,1)	(3/2,2,5/2)
Supplier 3	(1/2,2/3,1)	(2/5,1/2,2/3)	(1,1,1)

Normalized weight vector is $W_{LMC} = (0.39, 0.44, 0.16)$.

Applying Eq. (10) to calculate minimum degree of possibility:

$$d'(\text{Price}) = \min(0.89, 1) = 0.89, \tag{22}$$
$$d'(\text{Quality}) = \min(1, 1) = 1, \tag{23}$$
$$d'(\text{Flexibility}) = \min(0.57, 0.42) = 0.42. \tag{24}$$

Table 8. Fuzzy pairwise comparison matrix of alternatives with respect to product durability

	Supplier 1	Supplier 2	Supplier 3
Supplier 1	(1,1,1)	(1/2,1,3/2)	(2/7,1/3,2/5)
Supplier 2	(2/3,1,2)	(1,1,1)	(1,3/2,2)
Supplier 3	(5/2,3,7/2)	(1/2,2/3,1)	(1,1,1)

Normalized weight vector is $W_{PD} = (0.08, 0.39, 0.53)$.

Table 9. Fuzzy pairwise comparison matrix of alternatives with respect to supporting service quality

	Supplier 1	Supplier 2	Supplier 3
Supplier 1	(1,1,1)	(1,3/2,2)	(5/2,3,7/2)
Supplier 2	(1/2,2/3,1)	(1,1,1)	(2,5/2,3)
Supplier 3	(2/7,1/3,2/5)	(1/3,2/5,1/2)	(1,1,1)

Normalized weight vector is $W_{SSQ} = (0.62, 0.38, 0.00)$.

Table 10. Fuzzy pairwise comparison matrix of alternatives with respect to production capacity flexibility

	Supplier 1	Supplier 2	Supplier 3
Supplier 1	(1,1,1)	(1/2,2/3,1)	(2,5/2,3)
Supplier 2	(1,3/2,2)	(1,1,1)	(2/7,1/3,2/5)
Supplier 3	(1/3,2/5,1/2)	(5/2,3,7/2)	(1,1,1)

Normalized weight vector is $W_{PCF} = (0.41, 0.14, 0.45)$.

Table 11. Fuzzy pairwise comparison matrix of alternatives with respect to customization flexibility

	Supplier 1	Supplier 2	Supplier 3
Supplier 1	(1,1,1)	(2,5/2,3)	(1/2,1,3/2)
Supplier 2	(1/3,2/5,1/2)	(1,1,1)	(5/2,3,7/2)
Supplier 3	(2/3,1,2)	(2/7,1/3,2/5)	(1,1,1)

Normalized weight vector is $W_{CC} = (0.43, 0.42, 0.15)$.

The weight vector calculated as $W' = (0.89, 1, 0.42)$. After the normalization process, weight vector of main criteria price, quality, and flexibility was calculated $W_G = (0.39, 0.43, 0.18)$. The same procedure of calculation is used to determine priority weight of other pairwise comparison matrices (Tables 3, 4 and 5).

Therefore, supplier 1 is the best supplier for company among them because of high priority weight. The second best supplier for company is supplier 3 (Tables 6, 7, 8, 9, 10, 11 and 12).

Table 12. Summary of combined priority weights

Sub-criteria of price

Product price	Logistics management cost	Priority	Weight of alternatives
Weight	1	0	
Alternatives			
Supplier 1	0.51	0.39	0.51
Supplier 2	0.20	0.44	0.20
Supplier 3	0.29	0.16	0.29

Sub-criteria of quality

Product durability	Supporting service quality	Priority	Weight of alternatives
Weight	0.68	0.32	
Alternatives			
Supplier 1	0.08	0.62	0.25
Supplier 2	0.39	0.38	0.39
Supplier 3	0.53	0.00	0.36

Sub-criteria of flexibility

Production capacity flexibility	Customization flexibility	Priority	Weight of alternatives
Weight	1	0	
Alternatives			
Supplier 1	0.41	0.43	0.41
Supplier 2	0.14	0.42	0.14
Supplier 3	0.45	0.15	0.45

Goal of main criteria

Price	Quality	Flexibility	Priority	Weight of alternatives
Weight	0.39	0.43	0.18	
Alternatives				
Supplier 1	0.51	0.25	0.41	0.38
Supplier 2	0.20	0.39	0.14	0.28
Supplier 3	0.29	0.36	0.45	0.34

5 Conclusion

Suppler selection process is a broad and complex task for firms. In this paper, multi-criteria decision making approach fuzzy AHP has been used which includes both quantitative and qualitative criteria, to measure supplier performance. In this model, linguistic variables and triangular fuzzy numbers are used to overcome vagueness and uncertainty of decision makers. Multiple criteria helps decision makers to measure overall performance of supplier more efficiently. The advantage of this approach is that it takes less time and more accurate to solve supplier selection problem. Furthermore, fuzzy AHP model can be applied to any manufacturing company to choose best supplier. The application of this model is significantly effective and easily implement to choose best supplier.

Acknowledgements. The authors would like to give our great appreciation to editors and anonymous referees for their helpful and constructive comments and suggestions, which helped to improve this paper. This research is supported by the Youth Program of National Natural Science Foundation of China (Grant No. 71501137), the General Program of China Postdoctoral Science Foundation (Grant No. 2015M572480), the International Postdoctoral Exchange Fellowship Program of China Postdoctoral Council (Grant No. 20150028), the Project of Research Center for System Sciences and Enterprise Development (Grant No. XQ16B05), and Sichuan University (Grant No. SKQY201647).

References

1. Ahadian B, Abadi AGM, Chaboki RM (2012) A new DEA-GAHP method for supplier selection problem. Manage Sci Lett 2(7):2485–2492
2. Alencar LH, Almeida AT (2008) Multicriteria decision group model for the selection of suppliers. Pesquisa Operacional 28(2):321–337
3. Ayhan MB (2013) A fuzzy AHP approach for supplier selection problem: a case study in a gear motor company. Int J Manag Value Supply Chains 4(3):11–23
4. Boer LD, Labro E, Morlacchi P (2001) A review of methods supporting supplier selection. Eur J Purchasing Supply Manage 7(2):75–89
5. Chang DY (1992) Extent analysis and synthetic decision. Optim Tech Appl 1:352–355
6. Chou SY, Chang YH (2008) A decision support system for supplier selection based on a strategy-aligned fuzzy smart approach. Expert Syst Appl 34(4):2241–2253
7. Demirtas EA, Ustun O (2009) Analytic network process and multi-period goal programming integration in purchasing decisions. Comput Ind Eng 56(2):677–690
8. Dickson GW (1966) An analysis of vendor selection systems and decision. In: Materials science forum, pp 1377–1382
9. Dubois D, Henri P (1980) Systems of linear fuzzy constraints. Fuzzy Sets Syst 3(1):37–48
10. Gencer C, Grpinar D (2007) Analytic network process in supplier selection: a case study in an electronic firm. Appl Math Model 31(11):2475–2486
11. Ho W, Xu X, Dey PK (2010) Multi-criteria decision making approaches for supplier evaluation and selection: a literature review. Eur J Oper Res 202(1):16–24
12. Humphreys P, Mcivor R, Chan F (2003) Using case-based reasoning to evaluate supplier environmental management performance. Expert Syst Appl 25(2):141–153
13. Kahraman C, Tijen E, Buyukozkan G (2006) A fuzzy optimization model for QFD planning process using analytic network approach. Eur J Oper Res 171(2):390–411
14. Kumar M, Vrat P, Shankar R (2006) A fuzzy programming approach for vendor selection problem in a supply chain. Int J Prod Econ 101(2):273–285
15. Lee AHI (2009) A fuzzy supplier selection model with the consideration of benefits, opportunities, costs and risks. Expert Syst Appl 36(2):2879–2893
16. Li CC, Fun YP, Hung JS (1997) A new measure for supplier performance evaluation. IIE Trans 29(9):753–758
17. Pal O, Gupta AK, Garg R (2013) Supplier selection criteria and methods in supply chains: a review. Int J Soc Behav Educ Econ Bus Ind Eng 7(10):2667–2673
18. Prasenjit C, Poulomi M, Shankar C (2011) Supplier selection using compromise ranking and outranking methods. J Ind Eng Int 7(14):61–73
19. Roodhooft F, Konings J (1995) Vendor selection and evaluation an activity based costing approach. Eur J Oper Res 96(1):97–102

20. Saaty TL (1980) The analytic hierarchy process: planning, priority setting, resource allocation. McGraw-Hill, New York
21. Schramm F, Morais DC (2012) Decision support model for selecting and evaluating suppliers in the construction industry. Pesquisa Operacional 32(32):643–662
22. Shyur HJ, Shih HS (2006) A hybrid mcdm model for strategic vendor selection. Math Comput Model 44(7–8):749–761
23. Tam MCY, Tummala VMR (2001) An application of the AHP in vendor selection of a telecommunications system. Omega 29(2):171–182
24. Vahdani B, Zandieh M, Alemtabriz A (2008) Supplier selection by balancing and ranking method. J Appl Sci 8(19):3467–3472
25. Veerabathiran R, Srinath KA (2012) Application of the extent analysis method on fuzzy AHP. Int J Eng Sci Technol 4(7):649–655
26. Weber CA, Current JR, Benton WC (1991) Vendor selection criteria and methods. Eur J Oper Res 50(1):2–18
27. Wei S, Zhang J, Li Z (1997) A supplier-selecting system using a neural network. In: IEEE international conference on intelligent processing systems, vol 1, pp 468–471
28. Yahya S, Kingsman B (1999) Vendor rating for an entrepreneur development programme: a case study using the analytic hierarchy process method. J Oper Res Soc 50(9):916–930
29. Zadeh LA (1965) Fuzzy sets*. Inf Control 8(3):338–353

Cement Plant Site Selection Problem with Carbon Emission Trading Mechanism

Lurong Fan[1], Zongmin Li[2(✉)], and Pingwen Wu[3]

[1] Low Carbon Technology and Economy Research Center, Sichuan University, Chengdu 610064, People's Republic of China
[2] Business School, Sichuan University, Chengdu 610064, People's Republic of China
lizongmin@scu.edu.cn
[3] School of Economics, Sichuan University, Chengdu 610064, People's Republic of China

Abstract. Location problem is a fundamental task for the cement plant enterprises. This paper mainly investigates this problem with carbon emission trading mechanism. In practice, most of cement plant location problems are multi-objective and uncertain in nature. To enhance the general practice of cement plant location problem, this paper proposes a fuzzy random bi-level multi-objective model with carbon emission trading mechanism. Genetic algorithm is used to seek the approximate best solution of the model. A numerical example is also given to show the effectiveness of the proposed model.

Keywords: Cement plant location problem · Carbon emission trading mechanism · Genetic algorithm

1 Introduction

Cement is the most widely used essential construction material which is the main component of concrete, in addition, the cement industry is one of the most intensive energy consumptions [4,5]. Cement industry has been always among the largest CO_2 emission sources. Almost 5–7% of global CO_2 emissions are caused by cement plants, while 900 kg CO_2 is emitted to the atmosphere for producing one ton of cement [3].

Along with the phenomenon that carbon emission trading is accepted by many countries, the cement industry how to make product plan and select site to get the maximum profit and customer satisfaction under carbon dioxide trading mechanism, which are one of the most significant decision-making challenges and practical significance problem. As a consequence, the cement industry worldwide is facing growing challenges in conserving material and energy resources, as well as reducing its CO_2 emissions [7]. Because of the high carbon emission by cement industry, a series of researches have been constructed to attempt to reduce carbon emissions. Benhelal [2] presented new design of pyro-processing unit in a cement

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_59

factory and proved such novelties new process can significantly reduce 66% of CO_2 emission compared to the existing process. Madlool [6] found that carbon emission reduced by using energy efficiency measures in the cement industry.

However, there are few researches about cement plant site selection problem in the traditional article. Ataei [1] used a multi criteria decision-making method to rank alternative plant locations and found the best site. The problem of location of cement plant mainly has considered the cement industry. In fact, the customers purchase-choosing behavior should been taken into account in the cement plant location problem. It is often difficult to estimate cement sale revenues because these are decided by the customers. In turn, cement industry can also have an impact on the customers purchase-choosing behavior. Therefore, the cement plant location problem is formulated as a bi-level model with two decision-makers: cement industry and customers. Because the purchase-choosing behavior among the customers is uncertain, it is more reasonable to regard the customer's demands as a fuzzy-random variable. Therefore, the objective of this study aims to develop a fuzzy-random bi-level model for cement plant location problem with carbon emission trading mechanism.

The structure of the paper is as follows. Section 2 is the problem statement part. A fuzzy-random bi-level multi-objective model for location of cement plant with carbon emission trading mechanism is introduced in Sect. 3. And then a solution algorithm for the problem based on the KKT condition and GA (Genetic Algorithm) in Sect. 4. Section 5 gives a numerical example to illustrate the application of the model. Finally, some conclusions are presented in Sect. 6.

2 Problem Statement

The problem considered in this paper is a decision making problem between the cement industry and its consumers with carbon emission trading mechanism. For cement plant location problem, because each consumer is independent individual, cement plant site selection plan is not only determined by the cement industry, but also influenced by consumers. The cement industry choose some points to establish cement plants, decide the output and sale price, and after obtaining the cement price, the consumers make their decision in an consumer optimal manner to meet their demand and achieve the goal of minimal each cost at the same time. Therefore, the cement industry influences the consumers' purchase choosing behavior by adjusting the production plan and cement sale price, whereas each consumer tries to meet their demand to make the minimal cost based on decision made by the cement industry.

This paper considers the cement plant site selection problem as a leader-follower game where the cement industry is leader and the consumers who can freely chose the cement plant to purchase cement are the followers. In this situation, the cement plant site selection problem can be represented as a bi-level problem where the cement industry is the upper-level and the consumer is the lower-level (see Fig. 1).

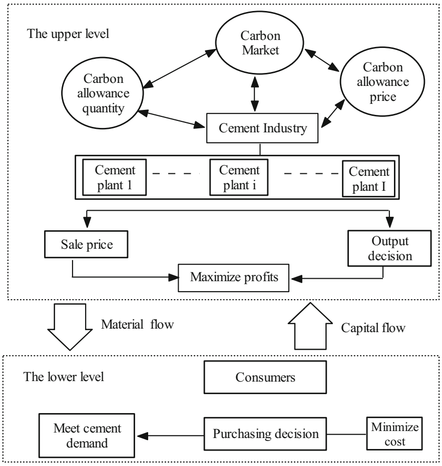


Fig. 1. The bi-level relationship between the cement plant and customers

3 Modeling Formulation

In this section, the relevant assumptions are first outlined, and then the bi-level model for the cement plant site selection problem under the carbon emission trading mechanism and considering fuzzy random is constructed. The mathematical description of the problem is given as follows.

3.1 Assumptions

To construct a model for the location problem of cement plant, the following assumptions are adopted:

(1) If customers buy cement, the cement plant must help the customer transport the cement without fee.
(2) The total alternative points is fixed.
(3) There are not any old cement plants before the new ones to be built.

3.2 Notations

The following mathematical notations are used to describe the cement plant location problem.

Index:

i: index of cement plants, where $i = 1, 2, \cdots, I$;
j: index of demand nodes, where $j = 1, 2, \cdots, J$;
n: index of cement species, where $k = 1, 2, \cdots, N$;

Certain parameters:

I: number of cement plants;
J: number of demand nodes;
N: number of cement species;
$t_j^{\max}$: the maximum tolerant transportation time of demand node j;
$t_j^{\min}$: the minimum transportation time of demand node j;
β_{n1}: the percentage of Ca in the cement n;
β_{n2}: the percentage of Mg in the cement n;
α_{i1}: the emission coefficient in product process at cement plant i;
α_{i2}: the emission coefficient in transportation process at cement plant i;
R: the initial allocated carbon allowance for cement company

Uncertain parameters:

$\tilde{\tilde{f}}_i$: the annual fixed costs for cement plant i;
$\tilde{\tilde{v}}_{i1}$: the unit variable costs for cement plant i in product process;
$\tilde{\tilde{v}}_{i2}$: the unit variable costs for cement plant i in transportation process;
$\tilde{\tilde{\eta}}$: the carbon emission quota price;
$\tilde{\tilde{a}}_{ijn}$: the unit transportation time of cement n between cement plant i and demand node j;
$\tilde{\tilde{D}}_{jn}$: the consumer j demand of cement n

Decision variables:

O_{in}: output decision of cement plant i about cement n;
u_i: the binary variable, if $O_{in} > 0$, $u_i = 1$, otherwise $u_i = 0$;
P_n: the sale price of cement n;
x_{ijn}: purchasing decision of demand node j

3.3 Upper-Level Model for Cement Plant Location Problem

It is not easy for the cement industry to make decision to get maximize profits and service satisfaction at the same time. Moreover, cement plant which is one of the major contributors to carbon emission, location problem is more meaningful to study with carbon emission mechanism.

(1) Objective Functions

In this paper, we consider cement plant which own two objectives, as the upper-level. In order to maximize economic profit, taking into account both improving the sales and saving the cost.

① Sale revenues

Cement industry revenues come from cement sales which related to the sale quantity and cement price. So the expected value of economic benefits is $\sum_{n=1}^{N} P_n \sum_{i=1}^{I} \sum_{j=1}^{J} x_{ijn}$.

② Carbon emission cost

The cement industry carbon emission comes from energy consumption for complete process and the raw material calcination process. The carbon emission by calcination process mainly from the reaction [3,8–10]:

$$CaCO_3 \rightarrow CaO + CO_2, \tag{1}$$

$$MgCO_3 \rightarrow MgO + CO_2. \tag{2}$$

From the reaction, we can calculate the carbon emission from the calcination process is: $\frac{11}{10}\sum_{n=1}^{N} O_{in}\beta_{n1} + \frac{11}{6}\sum_{n=1}^{N} O_{in}\beta_{n2}$. Approximately 50% of the total CO_2 emissions from the cement manufacturing process are released by the chemical processes and approximately 40% are released from burning fuel [10]. A minor contribution, 10%, is derived from electricity use or transport processes [11]. So the carbon emission from energy consumption includes produce processes and transport processes, i.e. $\alpha_{i1}\sum_{n=1}^{N} O_{in} + \alpha_{i2}\sum_{n=1}^{N}\sum_{j=1}^{J} x_{ijn}$. The carbon emission is limited by the government, the amount of cement industry's carbon emission would better not more than the limited standard. If initial carbon allowances allocated are sufficient, the carbon related cost is negative value, otherwise it indicates needing to buy carbon allowances. The carbon emission cost is as follows:

$$\left[\sum_{i=1}^{I}\left(\left(\frac{11}{10}\sum_{n=1}^{N} O_{ik}\beta_{n1} + \frac{11}{6}\sum_{n=1}^{N} O_{in}\beta_{n2}\right) + \alpha_{i1}\sum_{n=1}^{N} O_{in} + \alpha_{i2}\sum_{n=1}^{N}\sum_{j=1}^{J} x_{ijn}\right) - R\right] E_d(\tilde{\tilde{\eta}}). \tag{3}$$

③ Construction cost

The construction costs take into account fixed cost $E_d(\tilde{\tilde{f}}_i)$ and variable costs which are related to the cement quantity, $\sum_{n=1}^{N} O_{in}E_d(\tilde{\tilde{v}}_{i1}) + \sum_{n=1}^{N}\sum_{j=1}^{J} x_{ijn}E_d(\tilde{\tilde{v}}_{i2})$.

According to the aforementioned information, the equivalent for the total economic profits of cement industry is established as:

$$\begin{aligned} \max F = & \sum_{n=1}^{N} P_n \sum_{i=1}^{I}\sum_{j=1}^{J} u_i - E_d(\tilde{\tilde{\eta}}) \left[\sum_{i=1}^{I}\left(\left(\frac{11}{10}\sum_{n=1}^{N} O_{in}\beta_{n1} + \frac{11}{6}\sum_{n=1}^{N} O_{in}\beta_{n2}\right)\right.\right. \\ & \left.\left. + \alpha_{i1}\sum_{n=1}^{N} O_{in} + \alpha_{i2}\sum_{n=1}^{N}\sum_{j=1}^{J} x_{ijn}\right) - R\right] \\ & - \sum_{i=1}^{I} E_d(\tilde{\tilde{f}}_i) - \sum_{i=1}^{I}\left(\sum_{n=1}^{N} O_{in}E_d(\tilde{\tilde{v}}_{i1}) \sum_{n=1}^{N}\sum_{j=1}^{J} x_{ijn}E_d(\tilde{\tilde{v}}_{i2})\right). \end{aligned}$$

(2) Constraints

The cement industry will choose no more than M points from the total alternative points to built cement plant, so we have:

$$\sum_{i=1}^{I} u_i \leq M, \tag{4}$$

where u_i is decided by the cement plant output x_{ijn},

$$u_i = \begin{cases} 1, & x_{ijn} > 0 \\ 0, & x_{ijn} = 0. \end{cases} \tag{5}$$

It acquire the variable costs and fixed costs which will quickly be sank cost not more than G,

$$\sum_{i=1}^{I} \left(\sum_{n=1}^{N} O_{in} E_d(\tilde{\tilde{v}}_{i1}) + \sum_{n=1}^{N} \sum_{j=1}^{J} x_{ijn} E_d(\tilde{\tilde{v}}_{i2}) \right) + \sum_{i=1}^{I} E_d(\tilde{\tilde{f}}_i) \leq G. \tag{6}$$

3.4 Lower-Level Model for Cement Plant Location Problem

The lower-level problem represents the customers' choice behaviors and the demand of distributed among cement plants, that is to say each customer assigns his demand among the cement plants to minimize the total cost.

(1) Minimum Cost Objective

The lower-level objective is the minimum cost, it is decided by the cement price and purchase quantity, the customer's total cost is as follows:

$$\min L = \sum_{j=1}^{J} \sum_{n=1}^{N} P_n \sum_{i=1}^{I} x_{ijn}. \tag{7}$$

(2) Constraints

The purchase plan must satisfy the cement demand:

$$\sum_{i=1}^{I} x_{ijn} \geq E_d(\tilde{\tilde{D}}_{jn}). \tag{8}$$

The total purchase quantity of cement plant i is not more than the capacity of it:

$$\sum_{j=1}^{J} x_{ijn} \leq O_{in}. \tag{9}$$

3.5 Global Model

The lower-level customer choose cement plant and the amount of cement buying from it. The up-level cement industry choose some points from the fixed points to get the maximum profit and improve the service satisfactory; The customer try to influence the cement industry's decision to make their cost minimum. On the one hand, the lower cement price, the lower cost to the customer; On the other hand, the higher cement price, the higher profit to the cement industry. The location of cement plant can be represented as an game, they both can influence but can not control the other part.

The complete location problem involves the cement industry output decisions and sale price and the consumers purchase decision. Therefore, the global model includes bi-level objectives and constraints. In Noncooperative environments, we consider bi-level programming problems and it can be expressed as:

$$\begin{cases} \max\limits_{O_{in},P_n} F = \sum\limits_{n=1}^{N} P_n \sum\limits_{i=1}^{I}\sum\limits_{j=1}^{J} x_{ijn} - E_d(\tilde{\bar{\eta}}) \left[\sum\limits_{i=1}^{I} \left(\left(\frac{11}{10} \sum\limits_{n=1}^{N} O_{in}\beta_{n1} + \frac{11}{6} \sum\limits_{n=1}^{N} O_{in}\beta_{n2} \right) \right.\right. \\ \qquad \left.\left. + \alpha_{i1} \sum\limits_{n=1}^{N} O_{in} + \alpha_{i2} \sum\limits_{n=1}^{N}\sum\limits_{j=1}^{J} x_{ijn} \right) - R \right] - \sum\limits_{i=1}^{I} u_i E_d(\tilde{\bar{f}}_i) \\ \qquad - \sum\limits_{i=1}^{I} \left(\sum\limits_{n=1}^{N} O_{in} E_d(\tilde{\bar{v}}_{i1}) + \sum\limits_{n=1}^{N}\sum\limits_{j=1}^{J} x_{ijn} E_d(\tilde{\bar{v}}_{i2}) \right) \\ \text{s.t.} \begin{cases} \sum\limits_{i=1}^{I} u_i \le M \qquad i = 1, 2, \cdots, I \\ u_i = 0 \text{ or } 1, \qquad i = 1, 2, \cdots, I \\ \sum\limits_{i=1}^{I} \left(\sum\limits_{n=1}^{N} O_{in} E_d(\tilde{\bar{v}}_{i1}) + \sum\limits_{n=1}^{N}\sum\limits_{j=1}^{J} x_{ijn} E_d(\tilde{\bar{v}}_{i2}) \right) + \sum\limits_{i=1}^{I} E_d(\tilde{\bar{f}}_i) \le G \\ \text{where } x_{ijn} \text{ solve} \\ \min\limits_{x_{ijn}} L = \sum\limits_{j=1}^{J}\sum\limits_{n=1}^{N} P_n \sum\limits_{i=1}^{I} x_{ijn} \\ \text{s.t.} \begin{cases} \sum\limits_{i=1}^{I} x_{ijn} \ge E_d(\tilde{\bar{D}}_{jn}) \quad j = 1, 2, \cdots, J, \quad n = 1, 2, \cdots, N \\ \sum\limits_{j=1}^{J} x_{ijn} \le O_{in} \quad i = 1, 2, \cdots, I, \quad n = 1, 2, \cdots, N \\ x_{ijn} \ge 0 \quad i = 1, 2, \cdots, I, j = 1, 2, \cdots, J, \quad n = 1, 2, \cdots, N \end{cases} \end{cases} \end{cases} \tag{10}$$

4 Solution Approach

The proposed model (9) is a bi-level decision making problem, which reflects the interactive relationship between the cement industry and consumers. There are many algorithms available to solve bi-level programming. Therefore, the bi-level model (9) can be converted into a single level model with additional constrains by using KKT optimal conditions. So, the single level transformed from (10) by

KKT conditions is as follows:

$$
\left\{
\begin{aligned}
&\max_{O_{in},P_n} F = \sum_{n=1}^{N} P_n \sum_{i=1}^{I}\sum_{j=1}^{J} x_{ijn} - E_d(\tilde{\eta})\left[\sum_{i=1}^{I}\left(\left(\tfrac{11}{10}\sum_{n=1}^{N} O_{in}\beta_{n1} + \tfrac{11}{6}\sum_{n=1}^{N} O_{in}\beta_{n2}\right)\right.\right. \\
&\qquad \left.\left. +\alpha_{i1}\sum_{n=1}^{N} O_{in} + \alpha_{i2}\sum_{n=1}^{N}\sum_{j=1}^{J} x_{ijn}\right) - R\right] - \sum_{i=1}^{I} u_i E_d(\tilde{\tilde{f}}_i) \\
&\qquad -\sum_{i=1}^{I}\left(\sum_{n=1}^{N} O_{in} E_d(\tilde{\tilde{v}}_{i1}) + \sum_{n=1}^{N}\sum_{j=1}^{J} x_{ijn} E_d(\tilde{\tilde{v}}_{i2})\right) \\
&s.t. \begin{cases}
\sum_{i=1}^{I} u_i \le M \quad i = 1,2,\cdots,I \\
u_i = 0 \text{ or } 1, \quad i = 1,2,\cdots,I \\
\sum_{i=1}^{I}\left(\sum_{n=1}^{N} O_{in} E_d(\tilde{\tilde{v}}_{i1}) + \sum_{n=1}^{N}\sum_{j=1}^{J} x_{ijn} E_d(\tilde{\tilde{v}}_{i2})\right) + \sum_{i=1}^{I} E_d(\tilde{\tilde{f}}_i) \le G \\
\sum_{n=1}^{N} p_n - \lambda_{jn} - \varphi_{in} + \delta_{ijn} = 0, \quad i = 1,2,\cdots,I, \quad j = 1,2,\cdots,J \\
\lambda_{jn}\left(E_d(\tilde{\tilde{D}}_{jn}) - \sum_{i=1}^{I} x_{ijn}\right) = 0, \quad j = 1,2,\cdots,J, \quad n = 1,2,\cdots,N \\
\varphi_{in}\left(\sum_{j=1}^{J} x_{ijn} - O_{in}\right) = 0, \quad i = 1,2,\cdots,I, \quad n = 1,2,\cdots,N \\
\delta_{ijn}(x_{ijn}) = 0, \quad i = 1,2,\cdots,I, \quad j = 1,2,\cdots,J, n = 1,2,\cdots,N \\
\sum_{i=1}^{I}\sum_{n=1}^{N} x_{ijn} \ge E_d(\tilde{\tilde{D}}_{jn}), \quad j = 1,2,\cdots,J \\
\sum_{j=1}^{J} x_{ijn} \le O_{in}, \quad i = 1,2,\cdots,I, \quad n = 1,2,\cdots,N \\
x_{ijn} \ge 0, \quad i = 1,2,\cdots,I, j = 1,2,\cdots,J, \quad n = 1,2,\cdots,N \\
\varphi_{in} \ge 0, \quad i = 1,2,\cdots,I, \quad n = 1,2,\cdots,N \\
\delta_{ijn} \ge 0, \quad i = 1,2,\cdots,I, j = 1,2,\cdots,J, \quad n = 1,2,\cdots,N \\
\lambda_{jn} \ge 0, \quad j = 1,2,\cdots,J, \quad n = 1,2,\cdots,N
\end{cases}
\end{aligned}
\right.
\tag{11}
$$

The KKT conditions switched the cement sales from a game between cement industry and consumers to single decision making problem only by cement industry. The cement industry must try to control the sale price to get more market share. The consumers make purchase decisions by the cement price that reflects the competition.

5 Practical Application and Results

In this section, a case study of a practical cement plant selection problem in Conch cement company in China is introduced to demonstrate the effectiveness of the proposed model and optimization method.

5.1 Description of Case Problem

In this part, Conch cement company is taken as an application to explore the problem. Conch cement is a listed company with more than 80 subsidiaries in 17 province in China, and the parent company is located at Anhui province.

The average sales increased rate of Conch cement is 31.01% in last three years, ranking at 471 in total 1710 listed company and ranking at (9/29) in the construction industry. At the Copenhagen climate conference in 2009, the Chinese government pledged that China would reduce carbon emissions intensity (i.e. carbon emissions per unit GDP) by 40% $\sim$ 45% based on 2005 levels by 2020. In order to achieve this target, the government has proposed policy to control carbon emission. As one of the largest company in the construction industry, Conch cement must be more cautious when deals with decision problem like cement plant selection problem with limited carbon emission allowance.

5.2 Data Collection

Conch cement industry choose some points to build new cement plants which not more than 4 points in the possible 6 points. Every cement plants produce 2 different type cement. The Ca and Mg's percentage are different in the different cement which will influence the carbon emission from natural release, is shown in Table 1. The basic information for theses possible points, such as unit variable cost in production and transportation process, maximum output, emission coefficients in the production and transportation process are show in Tables 2 and 3. There are 3 different companies need buy some cements to build road,

Table 1. The percentage of Ca and Mg in the cement n

Cement n	Ca	Mg
$n = 1$	$\beta_{11} = 12.19\%$	$\beta_{12} = 20.31\%$
$n = 2$	$\beta_{21} = 23.16\%$	$\beta_{22} = 38.59\%$

Table 2. The basic information of the cement plants

Cement plant index i	Unit variable cost in produce process $\tilde{\tilde{v}}_{i1}$ (RMB/Tone)	Unit variable cost in transportation process $\tilde{\tilde{v}}_{i2}$ (RMB/Tone)	Maximum output l_i (10^4 Tone)
$i = 1$	(165, ξ_1, 179) $\xi_1 \sim N(172, 0.0258)$	(55, θ_1, 72) $\theta_1 \sim N(57.28, 0.0127)$	195
$i = 2$	(190, ξ_2 203) $\xi_2 \sim N(195, 0.0081)$	(69, θ_2, 81) $\theta_2 \sim N(70.34, 0.0215)$	168
$i = 3$	(230, ξ_3, 258) $\xi_3 \sim N(242, 0.0195)$	(96.2, θ_3, 101.7) $\theta_3 \sim N(98.5, 0.0097)$	175
$i = 4$	(262, ξ_4, 288.75) $\xi_4 \sim N(273, 0.0347)$	(79, θ_4, 105.2) $\theta_4 \sim N(83.2, 0.0438)$	210
$i = 5$	(181, ξ_5, 293) $\xi_5 \sim N(224.5, 0.0491)$	(73.9, θ_5, 87.15) $\theta_5 \sim N(76.5, 0.0278)$	151
$i = 6$	(176, ξ_6, 227.6) $\xi_6 \sim N(187, 0.0392)$	(82.05, θ_6, 97.63) $\theta_6 \sim N(86.42, 0.0365)$	220

Table 3. The basic information of the cement plants

Cement plant index i	Emission coefficient in produce process α_{i1} (kg/Tone)	Emission coefficient in transportation process α_{i2} (kg/Tone)	Annual fixed cost $\tilde{\tilde{u}}_i$ (10^8 RBM)
$i = 1$	302.12	90.24	$(1.0542,\ \varrho_1,\ 1.4518)$ $\varrho_1 \sim N(1.0589, 0.0653)$
$i = 2$	247.95	68.98	$(2.0786,\ \varrho_2,\ 2.4776)$ $\varrho_2 \sim N(2.2318, 0.1142)$
$i = 3$	444.44	103.65	$(0.8269,\ \varrho_3,\ 0.9287)$ $\varrho_3 \sim N(0.8675, 0.0625)$
$i = 4$	362.58	78.29	$(1.8841,\ \varrho_4,\ 2.0012)$ $\varrho_4 \sim N(1.9857, 0.0929)$
$i = 5$	233.92	57.48	$(2.5987,\ \varrho_5,\ 2.7014)$ $\varrho_5 \sim N(2.6013, 0.0471)$
$i = 6$	471.12	114.78	$(1.6638,\ \varrho_6,\ 1.8247)$ $\varrho_6 \sim N(1.6959, 0.0516)$

Table 4. The information of the consumers

Consumer index$j$	The consumers' demand of cement 1 $\tilde{\tilde{D}}_{1j}$ (Tone)	The consumers' demand of cement 2 $\tilde{\tilde{D}}_{2j}$ (Tone)
$j = 1$	$(48.4,\ \zeta_{11},\ 75.12)$ $\zeta_{11} \sim N(50.6, 0.0621)$	$(11.7,\ \zeta_{21},\ 24)$ $\zeta_{21} \sim N(12.4, 0.0315)$
$j = 2$	$(95.8,\ \zeta_{21},\ 119.9)$ $\zeta_{21} \sim N(105.37, 0.0323)$	$(26.6,\ \zeta_{22},\ 42.1)$ $\zeta_{22} \sim N(31.2, 0.0519)$
$j = 3$	$(146,\ \zeta_{31},\ 170.05)$ $\zeta_{31} \sim N(156, 0.0282)$	$(54.3,\ \zeta_{32},\ 89.7)$ $\zeta_{32} \sim N(76.7, 0.0724)$

airport, house, the relative information for these consumers, such as demand for the cement are show in Table 4. The data about *ca* and *mg* percentage in Table 1 are from the Conch cement industry. Most data in the Tables 2 and 3 were estimated based on expert consultations as well as the data in Table 4. The "allocation cap" was calculated to be 849.6×10^4 Tonnes, according to the total maximum output 944×10^4 Tonnes. The allowance price is $(20, \gamma, 60)$, where $\gamma \sim N(27.5, 10.36)$.

5.3 Results and Analysis

Almost 5–7% of global CO_2 emissions are caused by cement plants, therefore, the government allocated the allowance to the cement industry by the proportion. But to ensure the efficiency of the proved model, the carbon emission allowance R considering the following 3 scenarios.

Table 5. The purchase decision of the consumers

Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)	Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)	Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)
$j = 1$	$n = 1$	$x_{111} = 0$	$j = 2$	$n = 1$	$x_{211} = 9$	$j = 3$	$n = 1$	$x_{311} = 156$
		$x_{121} = 51$			$x_{221} = 6$			$x_{321} = 0$
		$x_{131} = 0$			$x_{231} = 0$			$x_{331} = 0$
		$x_{141} = 0$			$x_{241} = 0$			$x_{341} = 0$
		$x_{151} = 0$			$x_{251} = 0$			$x_{351} = 0$
		$x_{161} = 0$			$x_{261} = 90$			$x_{361} = 0$
	$n = 2$	$x_{112} = 14$		$n = 2$	$x_{212} = 30$		$n = 2$	$x_{312} = 0$
		$x_{122} = 0$			$x_{222} = 1$			$x_{322} = 77$
		$x_{132} = 0$			$x_{232} = 0$			$x_{332} = 0$
		$x_{142} = 0$			$x_{242} = 0$			$x_{342} = 0$
		$x_{152} = 0$			$x_{252} = 0$			$x_{352} = 0$
		$x_{162} = 12.4$			$x_{262} = 0.6$			$x_{362} = 0$

Table 6. The purchase decision of the consumers

Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)	Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)	Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)
$j = 1$	$n = 1$	$x_{111} = 53$	$j = 2$	$n = 1$	$x_{211} = 0$	$j = 3$	$n = 1$	$x_{311} = 25$
		$x_{121} = 0$			$x_{221} = 0$			$x_{321} = 0$
		$x_{131} = 0$			$x_{231} = 0$			$x_{331} = 0$
		$x_{141} = 0$			$x_{241} = 110$			$x_{341} = 0$
		$x_{151} = 0$			$x_{251} = 0$			$x_{351} = 0$
		$x_{161} = 0$			$x_{261} = 0$			$x_{361} = 130$
	$n = 2$	$x_{112} = 14$		$n = 2$	$x_{212} = 0$		$n = 2$	$x_{312} = 22$
		$x_{122} = 0$			$x_{222} = 0$			$x_{322} = 0$
		$x_{132} = 0$			$x_{232} = 0$			$x_{332} = 0$
		$x_{142} = 0$			$x_{242} = 33$			$x_{342} = 0$
		$x_{152} = 0$			$x_{252} = 0$			$x_{352} = 0$
		$x_{162} = 0$			$x_{262} = 0$			$x_{362} = 57$

(1) Scenario1: $R = 1 \times 10^6$ Tones

In this scenario, the allowance is controlled by the government very strict. Conch cement company should take action to decrease the carbon emission. The optimal cement price are 507.2 RBM/Tone and 700 RBM/Tone, and other optimal results are in Tables 5 and 7.

(2) Scenario2: $R = 4 \times 10^6$ Tones

In this scenario, the government allocated enough allowance to the cement company. Conch cement company can make full use of the resource. The optimal cement price are 447.29 RBM/Tone and 580.21 RBM/Tone, and other optimal results are in Tables 6 and 7.

(3) Scenario3: $R = 8 \times 10^6$ Tones

In this scenario, the government allocated more allowance to the cement company. Conch cement company can make full use of the resource and improve the cement production at the same time. The optimal cement price are 370.9 RBM/Tone and 550.21 RBM/Tone, and other optimal results are in Tables 7 and 8.

6 Policy Implications

The above results indicates that carbon emission allowance paly a important role in the cement site selection problem. When the government allocated low level allowance (i,e. 1×10^6 Tones) to cement company, it will control the output which just meet the consumers' demand, decrease the cost by limit the carbon emissions and make higher sale price; When the cement company gets medium level (i,e. 4×10^6 Tones), it will expand the scale suitably, increase the output and take lower sale price policy to improve the sale quantity; When the government allocated high level allowance (i,e. 8×10^6 Tones) to cement company, it will built more cement plants, take low price policy to arouse consumers extra demand to sale more cement.

As one of the most contributor to the carbon emission, the cement industry has different decisions with different allowance. On the one hand, building new cement plants and exploiting new market with high level allowance. The demand

Table 7. The output decision of cement plant

Cement Plant i	R = 1 $\times 10^6$ Tonnes		R = 4 $\times 10^6$ Tonnes		R = 8 $\times 10^6$ Tonnes	
	Cement 1 (10^4 Tones)	Cement 2 (10^4 Tones)	Cement 1 (10^4 Tones)	Cement 2 (10^4 Tones)	Cement 1 (10^4 Tones)	Cement 2 (10^4 Tones)
$i = 1$	$O_{11} = 165$	$O_{12} = 30$	$O_{11} = 145$	$O_{12} = 20$	$O_{11} = 131$	$O_{12} = 37$
$i = 2$	$O_{21} = 57$	$O_{22} = 78$	$O_{21} = 0$	$O_{22} = 0$	$O_{21} = 123$	$O_{22} = 28$
$i = 3$	$O_{31} = 0$	$O_{32} = 0$	$O_{31} = 0$	$O_{32} = 0$	$O_{31} = 118$	$O_{32} = 50$
$i = 4$	$O_{41} = 0$	$O_{42} = 0$	$O_{41} = 141$	$O_{42} = 40$	$O_{41} = 0$	$O_{42} = 0$
$i = 5$	$O_{51} = 0$	$O_{52} = 0$	$O_{51} = 0$	$O_{52} = 0$	$O_{51} = 0$	$O_{52} = 0$
$i = 6$	$O_{61} = 90$	$O_{62} = 13$	$O_{61} = 148$	$O_{62} = 62$	$O_{61} = 80$	$O_{62} = 100$

Table 8. The purchase decision of the consumers

Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)	Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)	Consumer j	Cement n	Purchase decision x_{ijn} (10^4 Tones)
$j=1$	$n=1$	$x_{111}=100$	$j=2$	$n=1$	$x_{211}=0$	$j=3$	$n=1$	$x_{311}=45$
		$x_{121}=0$			$x_{221}=123$			$x_{321}=0$
		$x_{131}=0$			$x_{231}=0$			$x_{331}=118$
		$x_{141}=0$			$x_{241}=0$			$x_{341}=0$
		$x_{151}=0$			$x_{251}=0$			$x_{351}=0$
		$x_{161}=0$			$x_{261}=0$			$x_{361}=50$
	$n=2$	$x_{112}=20$		$n=2$	$x_{212}=0$		$n=2$	$x_{312}=0$
		$x_{122}=0$			$x_{222}=28$			$x_{322}=0$
		$x_{132}=0$			$x_{232}=20$			$x_{332}=30$
		$x_{142}=0$			$x_{242}=0$			$x_{342}=0$
		$x_{152}=0$			$x_{252}=0$			$x_{352}=0$
		$x_{162}=0$			$x_{262}=0$			$x_{362}=80$

of cement will increased rapidly with the expanded of population scale, the cement company need to produce more cement to meet consumers' demand; On the other hand, the cement industry will control the output or do more research to exploit new technology to decrease carbon emission in the cement produce and transportation processes. Otherwise, if the cement company increase the output without new technology, it will face huge emission cost.

7 Conclusion

In this paper, a bi-level programming model with fuzzy random variables is proposed to deal with the cement plants site selection problem with carbon emission trading mechanism. Getting total profits maximum and making the total cost minimum are the cement company (i.e. the upper) and the consumers (i.e. the lower) targets, respectively. In contrast to the previous studies, the research takes the natural release and energy consumption which includes produce and transportation process as the total carbon emissions from the cement company. To solve the complex programming model, a KKT convert technique and a multi-cut points-based genetic algorithm are taken as a solution method. Finally, the results of the case study in Conch cement plants site selection problem are proved the efficiency of the optimization model and method.

Acknowledgements. We are thankful for financial support from the Research Center for Systems Science & Enterprise Development, Key Research of Social Sciences Base of Sichuan Province (Grant No. Xq15C01), the National Natural Science Foundation of China for Younger Scholars of China (Grant No. 71601134) and Project funded by China Postdoctoral Science Foundation.

References

1. Ataei M (2005) Multicriteria selection for an alumina-cement plant location in East Azerbaijan province of Iran. J S Afr Inst Min Metall 105(7):507–513
2. Benhelal E, Zahedi G, Hashim H (2012) A novel design for green and economical cement manufacturing. J Cleaner Prod 22(1):60–66
3. Benhelal E, Zahedi G et al (2013) Global strategies and potentials to curb co_2 emissions in cement industry. J Cleaner Prod 51(1):142–161
4. Kim SH (2015) Ventilation impairment of residents around a cement plant. Ann Occup Environ Med 27(1):3
5. Lamas WDQ, Palau JCF, Camargo JRD (2013) Waste materials co-processing in cement industry: ecological efficiency of waste reuse. Renew Sustain Energy Rev 19(1):200–207
6. Madlool NA, Saidur R et al (2013) An overview of energy savings measures for cement industries. Renew Sustain Energy Rev 19(1):18–29
7. Schneider M, Romer M et al (2011) Sustainable cement production present and future. Cem Concr Res 41(7):642–650
8. Stefanovi GM, Vučkovi G et al (2010) Co_2 reduction options in cement industry - The Novi Popovac case. Therm Sci 14(3):671–679
9. Strazza C, Borghi AD et al (2011) Resource productivity enhancement as means for promoting cleaner production: analysis of co-incineration in cement plants through a life cycle approach. J Cleaner Prod 19(14):1615–1621
10. Usón AA, López-Sabirón AM et al (2013) Uses of alternative fuels and raw materials in the cement industry as sustainable waste management options. Renew Sustain Energy Rev 23(4):242–260
11. WBCSD (2002) The Cement Sustainability Initiative: Our Agenda for Action. World Business Council for Sustainable Development (WBCSD), Switzerland

Representation and Analysis of Multi-dimensional Data to Support Decision Making in the Management

Shavkat Ayupov[1] and Abdulla Arifjanov[2(✉)]

[1] Institute of Mathematics, Academy of Sciences of the Republic of Uzbekistan, Tashkent, Uzbekistan
[2] Center for the Development of Software Products and Hardware-Software Complexes at the Tashkent University of Information Technologies, Tashkent, Uzbekistan
arifjanov@yandex.ru

Abstract. The paper considers questions related to the formal description of concepts used in the multivariate analysis of data, as well as to the set of operations on multidimensional cubes structuring the data. Discussed are questions in regard of the effectiveness of a class of multidimensional data models, and searching ways to solve the problem of "explosive" growth in the volume of data in a multidimensional OLAP-analysis.

Keywords: Data aggregation · Data hypercube · Data mining · Data selection · Data warehouse · Multidimensional representation of data · OLAP (On-Line Analytical Processing) · Operations in data manipulation

1 Introduction

An effective control of any production process, technical maintenance, or social configuration is not possible without taking timely, well-weighted adequate decisions. The complete work on such decisions requires the responsible teams and leaders to analysis large volumes of information, and as a rule, in conditions of rush and time limitations.

In the data analysis and decision support three basic concepts are used:

- The coding and structure of data base to ensure flexible navigation through the database, targeting generation of ad hoc queries and presentation of the results in the form of various reports. Their display should be based on the topological realization of algorithms for spatial analysis;
- The multi dimensional data manipulations to allow an easy organization of aggregated information from the storage in the self-saved form of a hypercubic model. This model would provide comfortable visualization and analysis

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_60

of data. Simultaneously, in the cells of a hypercube where numeric values are stored as aggregated indicators, the measurements to allow the organization of data in accordance to its chronology, geography and to other classifications, based on the reference storage;
- The search of dependences within the accumulated information to allow the use of algorithms for intellectual data analysis.

The concept of interactive multidimensional data analysis OLAP (On-Line Analytical Processing) was proposed in 1993, by Codd et al. [2], one of the ideologists of the relational databases. It was in the form of 12 basic rules that must be satisfied by the OLAP-systems, as well as the products providing the ability to perform online analytical processing. These rules can be considered as a theoretical basis for online analytical processing. In 1995, based on the Codd's et al. set of requirements, Nigel Pendse formulated the so-called Fast Analysis of Shared Multidimensional Information (FASMI) test, defining requirements for the applications of multidimensional data analysis.

OLAP-technology is an actual and essential topic of research. Its practical results are widely applicable. In this regard, the research analysis and presentation of multidimensional data have lately become a priority in the development of methods for designing information systems and systems for information analysis. The relevance of operational support systems for multidimensional data analysis has been around, and there are already many publications in this field. Despite a long history of research, materials on the formalized description of OLAP-systems are very few. In general, the available literature concerning this subject is devoted to the concepts, technology of OLAP-system, as well as to some practical issues. There can be found a lot of information on OLAP-systems, but virtually nowhere is written and said about how it works internally. Due to this reason, the solution of most problems in the construction of OLAP-systems is mainly carried out by the method of trials and errors. Currently, there is virtually no rigorous, unified mathematical foundation for systems of multidimensional data storage and analysis, such as that which exists for relational database management systems (DBMS). Little work is available on the formal description of the mechanism of aggregation in OLAP-cubes. Until now, there is no single terminology standard about data transmission, data query language, and about the formation of OLAP-cubes. In the existing publications it is mostly proposed multi-dimensional models describing the subject studied domains. Typically, these models were built for specific tasks. Moreover, their authors are using different terminology in description of various aspects of the subject area at different stages of design and data analysis. In order to be able to compare and analyze these works, we must have a unified system for classification of multidimensional models.

A rigorous mathematical description of the concept of multidimensional data and methods for constructing multivariate models provide several advantages. First and foremost, is to simplify the design process of multidimensional models, just as the normal form simplifies the design of relational models. Mathematical description of multidimensional operations can help to build simple and clear requests at the data analysis phase. Therefore, the development of a mathematical model for multi-dimensional representation of data and methods of analysis, is currently very topical problem. This article attempts to fill this gap.

Proceeding from the aforesaid, the purpose of this paper is: Firstly, to provide a formal description of concepts used in the multivariate analysis of data, as well as many operations on the multidimensional cube, by structuring the data; Secondary, to offer a study of the efficacy of a class of multidimensional data models, and to help in finding ways to solve the problem of "explosive" growth in the volume of data in multidimensional OLAP-analysis [1,2].

2 Basic Concepts of Multidimensional Data Model

Before proceeding to a formal description of a multidimensional data model, we adopt the basic concepts of multidimensional data model. Until now, there is no generally accepted point of view for the multidimensional models of data. Despite the absence of fundamental differences in existing views, however, there are some differences in terminology. We follow the terminology described in [1,3,5–7].

Basic concepts of multidimensional data models are:

- Data Hypercube is the entity of stored information;
- Dimension: a set of objects of one or several types organized in a hierarchical structure, which provides information context for a numeric indicator. Dimensions are usually visualized as the edges of a multidimensional cube;
- Facts: components of the measurements, tags (sequence of characteristics). These are objects, the totality of which forms the measurement. Facts are visualized as points or areas, located on the axes of the hypercube. These objects should be organized in a hierarchical structure so that the objects of the same type belong to only one level of hierarchy;
- Cells: atomic structures of the cube corresponding to particular values of some indicators. Cells in the visualization are inside the cube and are used to display the corresponding values of the indices;
- Indicator (Measure): a characteristic (usually numeric), which itself can be subject of analysis. An OLAP cube can have one or more indicators.

Data hypercube may contain one or more dimensions and is an ordered set of cells. Each cell is determined by the one and only one set of values of measurement–the facts. A cell can contain data - some measure or can be empty (Fig. 1).

In typical applications of OLAP-systems, it is usually used a "basic" attitude, which is called as the table of actual values (in short a fact table). Attributes of this attitude can be perceived as separate measurements of the multidimensional space in the form of a measurements table, or cube the original data (further called the hypercube), and each train - as a point in this space. At the intersection of the axes of measurement are located the data, quantitatively characterizing the analyzed facts C the measures. Data cube (or OLAP-cube) is different from the original data cube in the following aspects:

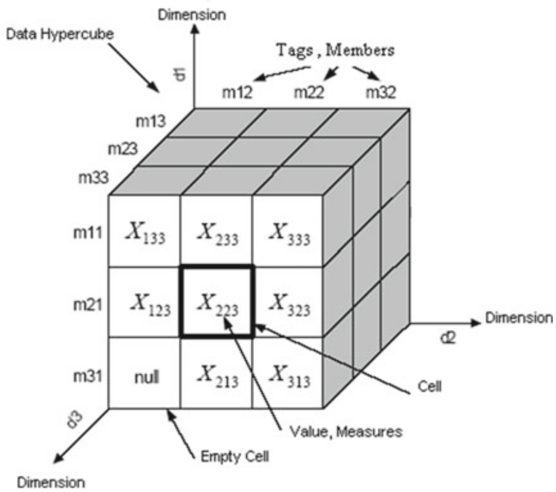


Fig. 1. An example of a data hypercube

- Along with the original data it contains information, aggregated over all subsets of measurements;
- Points in the data cube can be aggregated versions of the starting points of the original data. For example, the dimension of a "quarter" of the cube of initial data that determines the information about the progress for each quarter of a year may be reduced to the dimension of "half of a year".

3 A Mathematical Model of Multi-dimensional Representation and Data Analysis

OLAP systems represent software whose construction ideology is open towards analysis of large volume of structured information. Technology and application of class OLAP is focused on online analytical processing of multidimensional data sets, used to analyze the current and predict future activities of object management to managerial decision-making. OLAP technology simplifies the work with multi-purpose accumulated data, on the activities in decision making from the past with their extremely high volume. It also transfers sets of quantitative indicators into qualitative ones which are easy to understand. The system allows analysts, managers, and responsible personnel to form their own view on the data, using a fast, consistent, rapid access to diverse forms of information. These forms, based on primary data, allow the user to get a full picture of the activities which are subject of the control.

OLAP technology is an alternative to traditional methods of data analysis based on various systems implementation of SQL-queries to a relational database. OLAP systems play a crucial role in analysis and planning activities in the subjects of control. The analysis of data is mainly based on the requirements of decision makers to the information provided, traditionally existing individual characteristics of doing business in the objects of control, and decision-making mechanisms. From a user perspective the main difference between the OLAP systems and the Data Warehouse (DWH) systems is not technical, but in the subjected structure of the information.

The hallmark of OLAP systems is the representation of data in the form of multidimensional cubes (see e.g. [1,4]. All information about the business process is systematized, divided into categories, which, in turn, are converted to axes (measurements) of the cubes. Such an approach allows us to provide all the data in the form of two, three, or more dimensional sections of these cubes. The user can choose the sections on those or other categories, which allow combining, comparing and analyzing information in many possible directions. Moreover, each marginal direction can in turn be divided into sub-directions depending on the type of information.

Cube, this is a multidimensional database that stores two kinds of fields: measurement and evidence. Cube has a hierarchical structure which is optimized to be instantly used for immediate excerption of aggregated facts at the intersection of values (or components) of the measurements. The cube can contain pre-computed aggregates or provide special sampling data to calculate desirable units in real time. Its design provides instant performance and manipulations: data filtering, searching, changing the rules of aggregation, and more. In the sequel we consider OLAP features in detail.

Formal descriptions of terms used in the multivariate analysis of data, as well as many operations on multidimensional cubes structuring the data, we consider on the example of the data hyper-cube shown in Fig. 1.

Let consider first the algorithms for constructing a multidimensional data model to create a multidimensional database and data analysis using the mechanisms described in [1,3,5–7].

On Fig. 1 are depicted: $H(D, M)$ is the hypercube itself, i.e. plurality of cells; $V(H)$ is the set of measures; $D = \{d_{1,}d_{2,} \cdots, d_n\}$ is the set of measures; $M_{d_i} = \{m_{1_i}, m_{2_i}, \cdots, m_{k_i}\}$ is the set of measurements; - the set of labels pertinent to measurement d_i; $M = M_{d_1} \cup M_{d_1} \cup \cdots \cup M_{d_n}$ is the whole set of labels; $D' \subseteq D$ is a set of selected fixed dimensions; $M' \subseteq M$ is a set of selected fixed labels; $H'(D', M')|H' \subseteq H$ is a slice selected from the hypercube by the use of the sets D' and M'.

In order to gain access to the data, the user must specify one or more cells by selecting the measurement values (indicators), which correspond to the desired cells (the labels). This process of choosing some wished values of measurement is called fixing the facts, and the set of selected values of measurements - a set of fixed facts. The set of cells corresponding to the fixed labels and measurements is denoted as slice of the hypercube.

Hence, let $D = \{d_{1,}d_{2,} \cdots, d_n\}$ be the set of dimensions of the hypercube;

$M_{d_i} = \{m_{1_i}, m_{2_i}, \cdots, m_{k_i}\}$ is the set of labels for the measurement;
$M = M_{d_1} \cup M_{d_1} \cup \cdots \cup M_{d_n}$ is the set of facts contained in the hypercube;
$D' \subseteq D$ is a set of fixed dimensions;
$M' \subseteq M$ is a fixed array of facts.

The data hypercube consists of the array of all cells $H(D, M)$, corresponding to the sets D and M.

A subset of $H(D, M)$, corresponding to some subsets of fixed values D', M'. We denote as $H'(D', M')$.

Each cell h of the data hypercube $H(D, M)$ corresponds to only one possible set of facts measurement $M_h \subset M$. The cell can be empty (contains no data), or contains a value index - a measure.

The set of all measures included in the hypercube $H(D, M)$ is denoted by $V(H)$.

In a data hypercube one can perform either of the following operations (or manipulations) with data:

- operation "slicing";
- operation "rotation";
- operation "convolution and detail";
- operation "Data Aggregation";
- operation "sampling".

Consider these operations.

(1) The operation "slicing"
Any subset $H'(D', M')$ of the hypercube, obtained by fixing the labels from one or more dimensions, is called a slice. The operation of constructing the slice is carried out to obtain the desired subset of cells $H' \subset H$, and cutting the "unnecessary" values by successively fixing the needed tags. The slice can be interpreted as a two-dimensional array (a two-way table).

The label $m_{j_i} \in M$ specifies the hyper-plane section of a data hypercube, corresponding to an appropriate measurement $d_i \in D$. The set of fixed labels $M' \subseteq M$, thus defines a set of hyperplane intercepts (dissections) of the data hypercube corresponding to the set of fixed dimensions $D' \subseteq D$. The common intersection of these hyper-planes defines a set of cells $H'(D', M')$ is the slice of the data hypercube which is of interest to the user. The essence of the process of retrieving data from a data hypercube, thus consists in constructing a slice-a sub-hypercube of data $H'(D', M')$, by specifying the sets D', M'. Sometime, the result of this operation is called dice, and the operation itself is called dicing.

(2) The operation "rotation"
Any change in the order of presentation (imaging and visualization) of the measurements is called rotation. The rotations allow visualization of the data in forms, most comfortable for their perception.

(3) The operation "Data Aggregation" (Drill Up, consolidation, roll-up)
Data hypercube generally has hierarchies or formula-based relationships of data within each dimension. Aggregation involves computing all of these data relationships for one or more dimensions. While such relationships are normally summations, any type of computational relationship or formula might be defined. This means to get the values corresponding to the labels from a certain higher level $l, l1$ of the hierarchical dimension D based on the values from the previous lower hierarchical level $l - 1$.

Consider the hierarchical dimension D with L levels (Fig. 2). The primary data (events, facts) correspond to the lower level of the hierarchy ($l = 0$). Computation of aggregates is made in accordance with the applicable method of aggregation. In the case of summation, the values X_j^1 of the units at the level of the hierarchy $l = 1$ can be calculated by the formula $X_j^1 = \sum_{i=1}^{M_j} x_i^0$, where M_j is the number of values corresponding to the facts appearing subsidiaries to the label j.

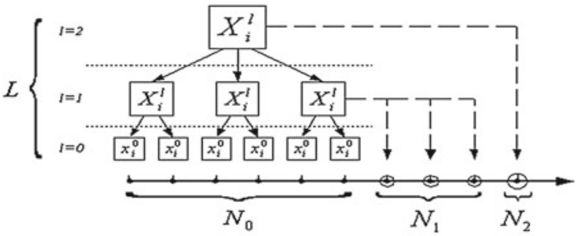


Fig. 2. Aggregation of data hypercube. One-dimensional representation

Summarizing, we obtain formulas for calculating the aggregated units according to the method of summation at other levels of hierarchy

$$X_j^l = \sum_{i=1}^{M_j} x_i^{l-1}, \quad l = 1, \cdots, L, \quad j = 1, \cdots, N_l.$$

The number of aggregated units in a data hypercube for a single measurement is $N_A = \sum_{i=1}^{L} N_i$.

Generalizing, for the case of an arbitrary number of dimensions D, for the number of aggregated units we obtain

$$N_A = \prod_{j=1}^{D} \prod_{i=1}^{L_j} N_i^j - \prod_{j=1}^{D} N_0^j,$$

where N_i^j is number of labels the i^{th} level of the hierarchy of measurements j, $l = 1, \cdots, D$, and L_j is the number of hierarchical levels of measurement j.

(4) The operation "convolution and detail" (Drill Down)
Drilling down is a specific analytical technique, where the user navigates among levels of data ranging from summarized data at higher level to more detailed data, located at lower level. This operation is carried out thanks to the hierarchical structure of the measurements. Readings (events, measurements) may be combined in a hierarchy, consisting of one or more levels. For example:

Day ⇒ Month ⇒ Quarter ⇒ Year;

Manager ⇒ unit ⇒ Region ⇒ Company ⇒ Country;

Model Car ⇒ The manufacturing plant ⇒ Country.

When going down to more and more detailed information the user is actually doing drill down.

(5) The operation "sampling"
This is a process of constructing a query to retrieve data from a data hypercube. It can be presented as a model of the network graph.

Consider the network graph $G(S, V)$ on Fig. 3, whose vertices correspond to the facts in the data hypercube $H(D, M)$. The set S of vertices consists of n subsets, or layers $S_i \subset S$. Among these are the "steps" of the user (a pattern) with which will be restricted the facts needed for the sample. The layer $S_i, i = 1, \cdots, n$ consists of specific vertices S_{j_i}, $j_i \in \{1, \cdots, |M|\}$, corresponding to facts that will be selected by the user on the i^{th} step.

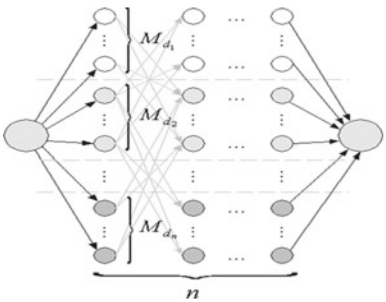


Fig. 3. Query construction

The set of edges of the graph, V, is a set of pairs of vertices $v^i_{jk} = \{s_{j_i}, s_{k_{i+1}}\}$. The edge $v^i_{jk} \in V$ of the graph, thus connecting the $j-th$ vertex S_{j_i} from the layer S_i with the $k-th$ vertex $S_{k_{i+1}}$ from the layer S_{i+1}.

The procedure for constructing a user request to retrieve data from a data hypercube $H(D, M)$ can be represented as a sequence of steps, at each of which the user captures a fact $m_{d_i} \in M$ from the measurement $\mathrm{d}_i \in D$. By fixing facts in desired dimensions, the user gets access to the cell corresponding to recorded and requested facts. The sequence of fixing a series of facts can be presented on the network graph $G(S, V)$ as a sequence of edges (paths) $w = \{v^0_{jk}; v^1_{jk}, \cdots, v^n_{jk}\}$, reflecting the process of formation of the user's query.

The initial vertex on the graph S_{0_0} corresponds to the beginning of the procedure of forming a user's query and to the absence of any fixed facts. In the first step ($n = 1$), the user fixes a fact $m_{d_i} \in M$ from the measurement $\mathrm{d}_i \in D$ of the data hypercube $H(D, M)$, and this corresponds to the vertex v_{j_i} on the layer S_1 of the graph $G(S, V)$. In the second step, the user captures a fact $m_{d_i} \in M \backslash M_{d_i}$ from the measurement $d_j \in D \backslash d_i$, since the measurement d_i has already been fixed. This procedure is performed consecutively n times, where n is the number of measurements needed to be fixed to obtain the required data slice (in general $n = |D|$).

The final vertex $S_{0_{n+1}}$ corresponds to the end of the procedure of forming a request and obtain access to the cell $h(w)$ of the data hypercube. Here w is the path selected by user's query on the graph $G(S, V)$.

In a diluted (scarce) data hypercube the requested cell $h(w)$, generally speaking, may be empty. Thus, time spent on the formation of a user query will be "wasted", because this particular query result is empty. The problem of optimizing the formation of user's queries by using step-by-step fixing facts measurement of the hypercube $H(D, M)$, is thus reduced to finding the set of paths W_{true} on a network graph $G(S, V)$, leading mainly to non-empty cells [7].

Unfortunately, the use of the described multidimensional data models and operations in the hypercube has shown poor performance on large volumes of data. For example, if the hypercube contains information about the sales in one year, and if it has only 3 measurements–Customers (250) Products (500) and Dates (365), we obtain a matrix of facts of the size $250 \times 500 \times 365 = 45625000$.

The total number of non-empty cells in the matrix can be practically only a few thousand. Moreover, the more the number of dimensions, the more tenuous will be the matrix. Therefore, to work with such matrix it is preferable to create and use special methods of processing data in sparse matrices. To solve this problem, it is possible by preliminary "cleaning" the data before to use them to build the cubes. But this approach is not always applicable. Another drawback is that the choice of a higher level of minuteness in the creation of a hypercube can greatly increase the size of the multidimensional database. Because of these, and some other additional reasons, commercially available multidimensional database management systems are not able to handle large amounts of data. It is advisable to use this multivariate model, when the size of the database is small and has a homogeneous in time set of measurements.

In contrast to the multidimensional databases, the relational management systems of databases are capable to store huge amounts of data. But they lose their advantages when compared in the speed of execution of analytical queries. When using relational database management systems, the principles for data storage are organized in a special way. The most commonly used so-called radial (or "star") pattern. In this scheme, two types of tables are used: (1) Fact table and (2) Several reference tables (dimension tables). The fact table usually contains data, the most intensively used for analysis of data in the cube. If we draw an analogy with the multivariate model, a record in the factual table corresponds to a cell in the hypercube. In the reference tables are listed all the possible values of each one dimension of the hypercube. Each dimension is described by its own lookup table. Fact table is indexed on a complex key that is built from individual keys help reference tables. It provides a link-up table with the factual by suitable sets of key attributes.

To shorten the response time in an analytical system one can use some special tools. In the composition of powerful relational database management systems there are typically included query optimizers. When creating a data warehouse based on relational database management systems, their presence is of particular importance. Optimizers analyze the request and determine the best position, in regard to some criterion, of the sequence of operations to access the database for this specific request implementation. For example, in this way one can minimize the number of physical disk accesses for the query. The query optimizers use sophisticated algorithms for statistical processing, which operate on the number of entries in the tables, ranges of keys, etc.

Each of the models described above has both, advantages and disadvantages. Multivariate model allows implementation of fast analysis of data, but cannot store large amounts of information. Relational model, by contrast, has virtually no limit on the amount of accumulated data, but such database does not provide the desirable speed in the run of analytical queries, as the multidimensional database.

4 Ways to Solve the "Explosive Volume Growth" Problem in Multidimensional OLAP

Data warehouses and OLAP analytical system are designed for analysis and synthesis of detailed data. So, they can include a full range of detailed data contained in database sources, and may increase the amount of data as a result of denormalization and duplication of detailed data. In addition, the amount of data can grow when aggregated information is added to the detailed data. This explains the rapid growth of data in the OLAP structures, as compared with the rate of increase of the volume in source databases. Often to describe this phenomenon, the experts are using the term "explosive" nature of increase of data in hypercube.

Accelerated growth of data volumes dramatically reduces the performance of multivariate analysis and limits its scalability. We have investigated various strategies to improve query performance in relational data warehouses and analytical OLAP-systems. We tried to achieve the rapid growth of the processed data and typical compact solutions without built-in mechanisms to improve system's performance and scalability. Considered were the following strategies:

- Distributing the data according to its relevance of use;
- Partitioning the storage fact table and the hypercube structure into sections and into sub-cubes of smaller size;
- Organization of parallel query processing.

The results show that the largest integrated effect on improved performance and scalability provides a combined strategy. It combines the strategies of partitioning and parallel processing of sections. While the available storage capacity increases in proportion to the number of sections in the fact table, the query performance increases by several times. It was found that the specificity of the multivariate analysis cannot directly transfer the methods for partitioning fact tables on the hypercube. This is due to the difference in the processes of indexing data in hypercube, and these in relational tables.

As noted above, OLAP-technology has emerged as an alternative to the traditional methods of data analysis based on limited and SQL-queries to a relational database. The main advantage of OLAP lies in the wide possibility of forming ad-hoc queries for analytical databases. However, our studies show that from a theoretical point of view, unlimited opportunity to form ad hoc queries to OLAP hypercube when working with the above model, under certain conditions can lead to complex problems, associated with ensuring the integrity of the multidimensional data. A similar problem arises, for example, when using the methods of the additive decomposition of a hyper-cube's structure into sub-cubes of smaller size. In such cases an ad hoc query, being addressed to the complete hypercube (as a whole), first processes the sub-cube, and the results are then consolidated into the overall result. With ad-hoc queries indexing hypercube becomes arbitrary, varies from request to request, and may contribute to errors in the procedure of consolidation in the settlement of the results from sub-cubes. We have investigated the issue and presented a new method for its solution. It eliminates data corruption, despite the non-additive indexing of additively decomposed hypercubes. The conditions for indexing data in sub-cubes were analyzed. An assertion was stated, that allows to identify any subset of undistorted data sub-set X_{sec} and the potentially faulty data subsets X'sec at the intersections of indices in the following form

$$X_{\text{sec}} \equiv X \backslash X'_{\text{sec}} = X \backslash \overset{n-1}{\underset{i=1}{U}} \big(\underset{i<j\leq n}{U} (X_i I X_j) \big),$$

here n is the number of decomposed sub-cubes, and X is a notation for the original hypercube:

$$X = X_{\text{sec}} \cup X'_{\text{sec}};\ (X_{\text{sec}} | X'_{\text{sec}} = \emptyset).$$

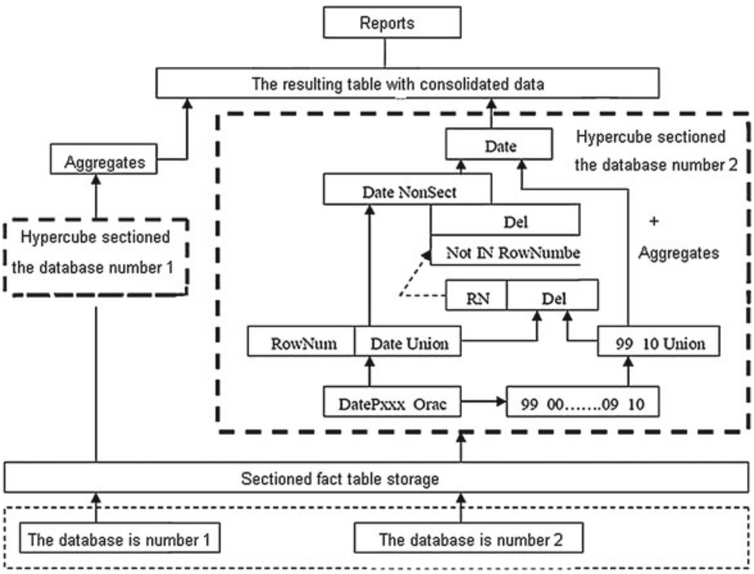


Fig. 4. The structure of the algorithm of multidimensional data aggregation

In accordance with the obtained by us conditions for the algorithm of multi-dimensional data aggregation, the overall structure is shown in Fig. 4.

The developed algorithm is tested in the solution of specific applications. It allows the use of a convenient way for additive decomposition of hyper-cube's structure in cases of ad hoc requests from users, resulting in non-additive indexing of the data in the hypercube.

5 Example

As an example, consider the formation of OLAP-cube based on existing relational fact table sales from managers, M1, M2 and M3 cars Nexia, Tico and Damas in the years of 2007, 2008 *and* 2009. To do this, it is needed to select the values of all measures and carry out the aggregation of all the obtained values, and record it in the cube. Consider the application of the obtained models on the example of a cube with 3 dimensions (Fig. 5).

Here measurements are: the model made of cars (products, a categorical variable), the sale person (manager, another structural attribute), and the year of production (time dimension), i.e. this is the triplet C, M, T;

Facts are: Model car (Nexia, Tico, Damas), labeled as (N, T, D); Managers (Manager 1, Manager 2, Manager 3) labeled as $(M1, M2, M3)$; The time (year) recorded $(2007, 2008, 2009)$ labeled as $(Y4, Y5, Y6)$;

Measures are: the volume of sales $(7, 7, 4, etc.)$.

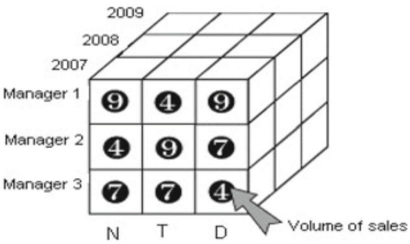


Fig. 5. A cube with three dimensions

In our example we have:

The set of dimensions of the hypercube is $D = C, M, T$;

The set of labels for the measurement d_i: $M_C = \{N, T, D\}$, $M_M = \{M1, M2, M3\}$ and $M_T = \{Y4, Y5, Y6\}$.

In this example each value in the field of sales is uniquely determined by a combination of fields: Model car, Manager, Year of sale. The indicator is the sales volume, located in the cell of the intersection of specified dimension of each kind.

Consider the application of operations to manipulate dimensions:

(1) Formation of a "Slice"
If we limit the value "dimensions of a model car" = "Nexia GL", then we obtain a subset of the hypercube (in our case - two-dimensional table) containing information about the history of sales of this model by different managers in different years. This is the content on the left wall of the cube on Fig. 5.

(2) Operation "Rotation"
This operation provides the ability to visualize data in a manner most comfortable for its perception. Rotating consists of swapping the rows and columns. For example, by changing the Measures of Products down on front X-axis by the Time measure across the side Y-axis into a report, an entirely new vision of the data hypercube on Fig. 4 will be obtained. This will be equivalent to a rotation in 90° counter clock wise along the Z-axis (the managers).

(3) Operation "Aggregation"
To get the sales for each car model and for each year, we need to accumulate the numbers in the cell of each column for all the managers. To get the total sales of each manager throughout the years we need to add his/her results for each car model for all years. This is aggregation.

From user's perspective, in a larger dimension hierarchical model Division, Region, Company, Country is exactly the same kind of measurements as the manager is. But each of these added dimensions would correspond to a new, higher level of aggregated values of the index "volume of sales". In the process of analysis the user works not only with a variety of data slices and performs their rotation, but also moves from detailed data to aggregated, i.e. applies the operation Aggregation. For example, after noticing how successfully Manager 1 sold the models of "Nexia" and "Tico" in 2007, his supervisor may want to see how looks its relationship with the sales of these models at the unit level, where Manager 1 works. And after that the supervisor may want to obtain a similar certificate for the Region or firm.

(4) Operation "Detailization"
The transition from aggregated to more detailed data is the operation "Detailization". For example, beginning the analysis at the level of total sales through all the years, the supervisor may wish to obtain more precise information about specific local sale in a year and by a $Manager$.

(5) The "Sampling"
To construct a query to retrieve data from the data hypercube presented on Fig. 4 a user need to select a sequence of addresses (vertices of an associated graph) of cells whose content are required in the sample. For instance, the choice $(N, M2, Y5) \Rightarrow$ (T, M3, Y4) $\Rightarrow (T, M1, Y6)$ will retrieve the content of the corresponding cells on the display in this order.

6 Conclusions

It is shown that the use of the conventional multidimensional data models and operations in the hypercube does not provide the desired speed performance on large volumes of data. This problem can be solved by pre-treatment of data before to use them to build the data cubes. Another drawback is that the choice of a high level of detail in the creation of a hypercube can significantly increase the size of a multidimensional database. Because of these and some other reasons, commercially available multidimensional database management systems are not able to handle large amounts of data. It is advisable to use a standard (conventional) multivariate model, if the size of the database is small and has a homogeneous set of measurements. Largest integrated effect of improved performance and scalability provides a combined strategy that combines partitioning and parallel processing sections. Then the available storage capacity increases in proportion to the number of sections in the fact table, and query performance increases by several times.

It was found that the specificity of the multivariate analysis cannot immediately transfer the methods for partitioning the fact table on the hypercube. This is due to the difference in the processes of indexing data for hyper-cubes, and for relational tables.

References

1. Arifjanov AS (2008) Algorithms for multidimensional reporting and analysis of data. Uzbek J Probl Inf Energ 8(1):65–72
2. Arifjanov AS (2014) Some questions in mathematical modeling of multidimensional data reporting and analysis. In: Book of abstracts. Flint international statistical conference “FLINT:1 City Hundred Years under Variability”
3. Ayupov S, Arifjanov A (2017) Information–analytical technologies of decision support in management of power systems. Springer, Singapore
4. Codd EF, Codd SB, Salley CT (1993) Providing OLAP (on-line analytical processing) to user-analysts: an it mandate, vol 32. E.F. Codd & Associates, pp 12–20
5. Hrustalev EM (2013) Aggregation of data in olap-cubes. http://www.olap.ru
6. Sakharov AA (2017) The concept of building and implementing information systems, with a focus on data analysis. http://olap.ru/basic/saharov.asp (in Russian)
7. Zabotnev MS (2017) Methods of presenting information in a sparse data hypercube. http://www.olap.ru

Fuzzy Logic Applied to SCADA Systems

Tahar Benmessaoud[1,2], Alberto Pliego Marugán[2], Kamal Mohammedi[1], and Fausto Pedro García Márquez[2](✉)

[1] Laboratory of Energetic-Mechanic and Engineering, FSI, M'hamed Bougara University of Boumerdes, Boumerdès, Algeria
[2] Ingenium Research Group, Universidad Castilla-La Mancha, 13071 Ciudad Real, Spain
FaustoPedro.Garcia@uclm.es

Abstract. This article focuses on the monitoring of a wind farm in real time based on big data collected by Supervisory Control and Data Acquisition (SCADA) system. The decision-making of the type of maintenance to be applied can be insured by SCADA system. This system generates alarms based on the collected data. False alarms cause false interventions by the maintenance team resulting in loss of production and costs. The reduction of these false alarms makes it possible to contribute better to the management of the maintenance of the wind farm. In this paper, we propose a new approach for the identification of alarms by Fuzzy Logic based on the data collected by the SCADA system. The alarms generated in this case can be divided into two categories: orange alarms corresponding to faults requiring the intervention of preventive maintenance and red alarms corresponding to critical states that can cause system failures.

Keywords: Wind farm · Monitoring · SCADA · Alarms · Fuzzy logic approach

1 Introduction

The generation of electricity through wind energy systems is demanded by the industry in order to increase the competitiveness of the business. The maintenance of these systems becomes very complex because the stochastic city of the random loads can cause catastrophic failures [3,12]. The optimization of the maintenance planning is a crucial factor for the efficiency of the wind farms [15]. With this purpose, wind turbines (WT) can be monitored by SCADA or condition monitoring systems to detect failures [13,14], risks and to take necessary actions online [8]. Several studies have proved the efficiency of this systems in the WTs [9] and other types of industries [7,13,19].

False alarms generated by SCADA system are a important problem because they cause unnecessary stops, false interventions by the maintenance team, loss of production and, consequently, extra costs [11,17,18]. Some studies aim to

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_61

eliminate false alarms of WTs [5]. Chen et al. [5] present a study on the treatment of SCADA using an artificial neural network 'ANN'. Qiu et al. proposes two methods (sequential and probabilistic) for the analysis of alarms for two large onshore wind farms considering the defects of pitch and converter systems [20].

This paper proposes a new approach based on fuzzy logic for controlling wind farms considering the big data collected by SCADA system, and give a complementary response to strengthen the response of the SCADA. Figure 1 shows the general structure of a fuzzy system [1].

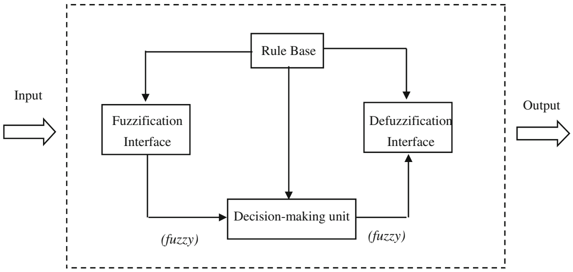


Fig. 1. Basic configuration of a fuzzy logic system

In the fuzzification part, any numeric value of the input data is converted into a linguistic value between 0 and 1. The fuzzy system requires functions for each input data:

- Input (1): Variable (1). Subsets: low, average and maximum.
- Input (2): Variable (2). Subsets: low, average and maximum.
- Input (3): Variable (3). Subsets: low, average and maximum.
- Input (n): Variable (n). Subsets: low, average and maximum.

Fuzzy inference is a method that interprets the values in the input vector and, based on fuzzy rules, assigns values to the output vector [16]. With this purpose, it is necessary to stablish the rules that will perform the defuzzification process to provide the output. The Fuzzy rules (IF antecedent THEN consequent) in expert system are usually is following [10]:

IF Var(1) is $A11$ and/or Var(2) is $A21, \cdots$ THEN y is $B1$
else
IF Var(1) is $A12$ and/or Var(2) is $A22, \cdots$ THEN y is $B2$
else

IF Var(1) is $A1n$ and/or Var(2) is $A2n, \cdots$ THEN y is Bn
where $\mathrm{Var}(1), \mathrm{Var}(2), \cdots, \mathrm{Var}(n)$ are the fuzzy input (antecedent) variables, y is a single output (consequent) variable, and $A11, \cdots, A1n$ are the fuzzy sets [5]. Generally, there are n input variables consisting of 3 fuzzy linguistic variables. Therefore, a total of 3^n rules will provide all possible combinations of the input variables.

Defuzzification is the process of producing a quantifiable result given fuzzy sets and corresponding membership degrees. There are many types of defuzzification methods, usually maximum membership and centroid techniques are used [22].

2 Methodology Proposed

In this paper, a novel methodology is presented to make an original treatment of the SCADA dataset. The flowchart of the methodology proposed is shown by Fig. 2.

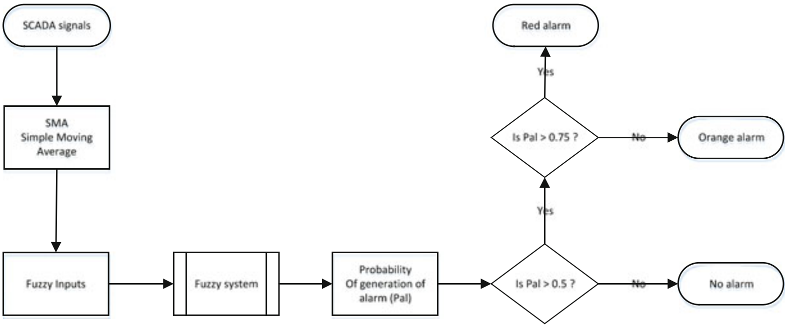


Fig. 2. Alarm identification flowchart

In the diagram adopted for the identification of alarms, it is integrated the real data collected by SCADA system with the fuzzy system. With this purpose, the SCADA data will be preprocessed before of being inputted in the fuzzy system. This preprocess starts with the evaluation of the simple moving average (SMA) of the SCADA signals. Then, the difference between the signals and the SMA is calculated. These differences will be used as input vectors for the fuzzy system.

Once the new variables have been obtained, the next step is to establish some control laws for them. The new variables will be analyzed statistically to obtain two thresholds. These thresholds will separate the graph into three different regions. Figure 3 shows the definition of the thresholds and the regions (good, acceptable, unacceptable) considered

This classification allows for determining the member functions of the inputs of the fuzzy system and generate the fuzzy system. The control of each physical parameter collected by SCADA can be represented by the difference between the values of all the peaks and the mean value. The fuzzy inference system is based on different rules to generate the occurrence probabilities of the alarms in the output.

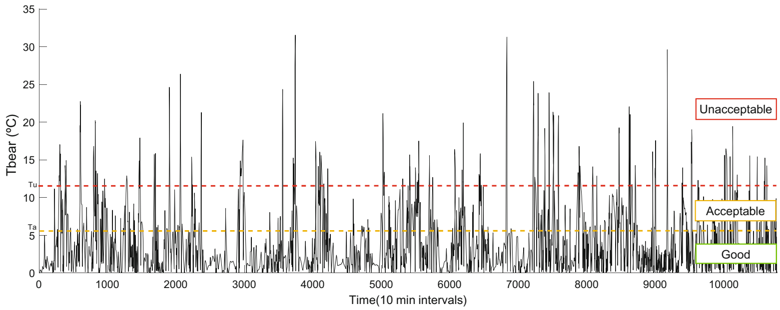


Fig. 3. Classification for member functions

3 Real Case Study

The data considered for this real case study is obtained from the European Project entitled OPTIMUS [4]. The database used in this paper come from a SCADA system that provides different measures every 10 min in the period from 01/01/2015 to 28/03/2015. The system measure 37 physical variables, but in this paper only 4 parameters have been considered to simplify the example. These parameters correspond to the speed of main shaft (Vel), vibration at the main shaft (Vibr), oil temperature (Toil) and bearing temperature (Tbear). Figure 4 shows these parameters.

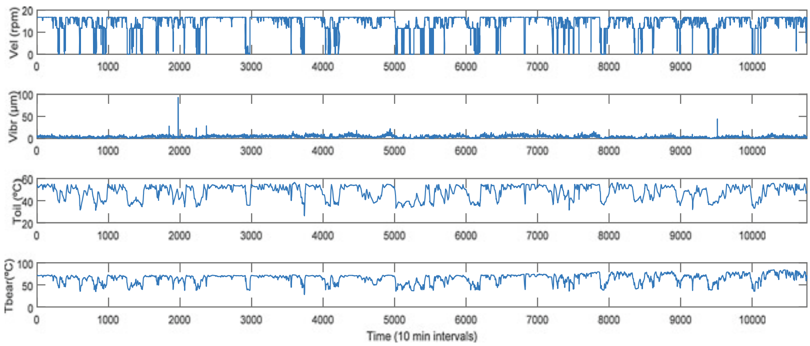


Fig. 4. Evolution of real data in time

Figure 5 shows the preprocessing of the variable Tbear before of being inputed in the fuzzy system. The blue line corresponds to the “Temperature of bearing (Tbear)”. The red line evaluates the moving average (MA) calculated for a period of 2 h. The black line represents the absolute value of the difference blue and the red line. This new variable will be named “Difference of Tbear” and noted (DTbear).

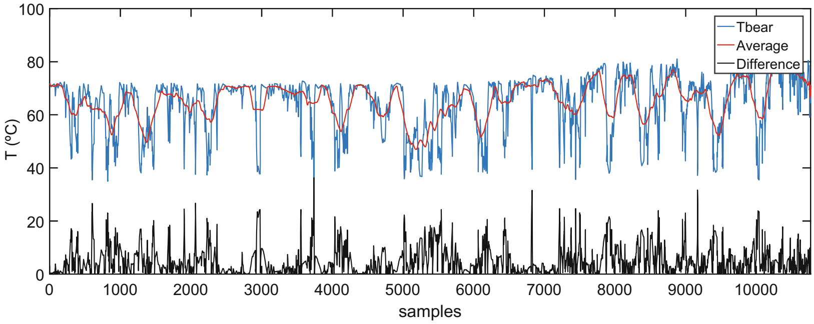


Fig. 5. Preprocessing of SCADA signals

The same procedure will be applied to the rest of variables. The new variables will be obtained by the following equations:

$$\begin{aligned}
\mathrm{Dvel} &= \mid \mathrm{Vel} - \mathrm{SMA}(\mathrm{Vel}) \mid, \\
\mathrm{DVibr} &= \mid \mathrm{Vibr} - \mathrm{SMA}(\mathrm{Vibr}) \mid, \\
\mathrm{DToil} &= \mid \mathrm{Toil} - \mathrm{SMA}(\mathrm{Toil}) \mid, \\
\mathrm{DTbear} &= \mid \mathrm{Tbear} - \mathrm{SMA}(\mathrm{Tbear}) \mid .
\end{aligned}$$

Consequently, Fig. 4 will be transformed to Fig. 6, considering the new variables.

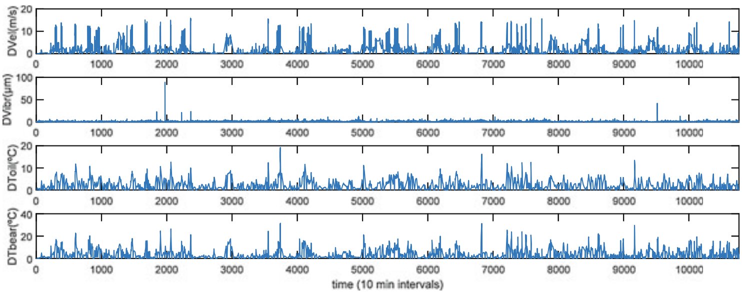


Fig. 6. Evolution of real data in time

For stablishing the different thresholds pointed in Fig. 3, a maximum value have been assigned to each input considering several norms and standards [2,6,12,21]. These values correspond to:

- max(Rotor velocity) = 20.2 rpm.
- max(Vibration) = 115 μm, max(Temperature of oil) = 95 °C.
- max(Temperature of bearing) = 110 °C.

Thus, some of the rules defined in this example are:

- IF DVel is good and DVibr is good and DToil is good and DBear is good THEN the output is green.
- IF DVel is good and DVibr is acceptable and DToil is unacceptable and DBear is Good THEN the output is red.
- IF DVel is unacceptable and DVibr is good and DToil is unacceptable and DBear is unacceptable THEN the output is red.
- If DVel or DVibr or DToil or DBear is unacceptable then the output is red.

More rules of the Fuzzy system are presented graphically in the Fig. 7:

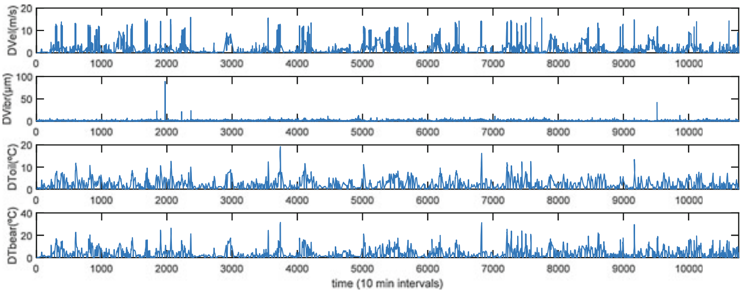


Fig. 7. Rule based Fuzzy Inference System for generation of probabilistic alarms

The red line of each variable shows the value of the variable. The displacements to the left or the right of these lines generates a new position of the output, and therefore, a new probability of alarm. Once, the fuzzy system is defined, it is possible to represent surfaces that explain the behavior of the system under different conditions. For example, Fig. 8 represents the probability of alarm depending on DTbear and DVibr.

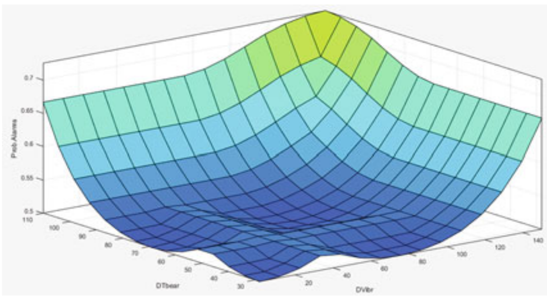


Fig. 8. Surface view of alarms probability with respect to DTbear and DVibr

It is noted that the probability varies proportionally with the temperature of the bearings and the vibration. For values of DTbear and DVibr respectively less than certain value, the probability has very minimal values, in opposite, the probability of alarm increase exponentially when these values are exceeded.

The system built has been proved by performing a simulation with the SCADA data. The results of the simulations are shown in Fig. 9. The system provides a certain probability of alarm regarding to the inputs.

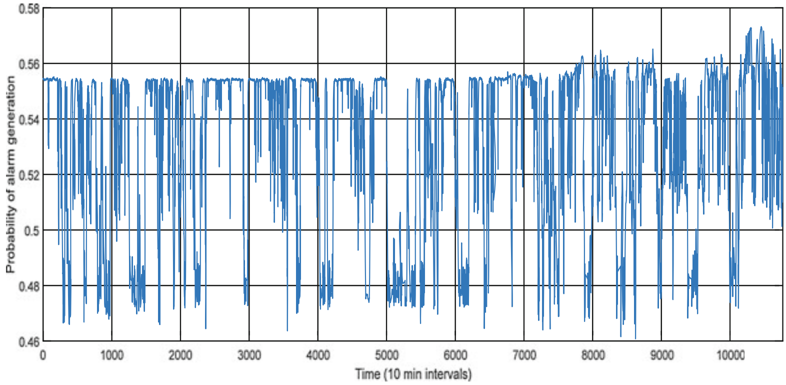


Fig. 9. Probability of alarms generation

A total of 10768 inputs have been analyzed through the created fuzzy system. The outcomes are:

- 10407 normal measures. This represent the 96.7% of the total data.
- 361 orange alarms. This represent a 3.3% of the total data.
- No red alarms. In the period studied there is not any value exceeding the limits of critical alarm.

This method allows to transform the data collected by the SCADA system into probability of alarms. It can be a useful information for complementing the decision making. The methodology can aid to reduce false alarms because when a critical alarm arises from the SCADA system, the response of the fuzzy system can reinforce that alarm.

4 Conclusions

In this paper, a new methodology based on fuzzy logic is proposed in order to analyze the SCADA data and provide an alternative decision support. The main purpose is to process the signals collected by the SCADA system from a different perspective. These signals are converted into new variables to be inputted in the fuzzy system. A fuzzy system has been created using several standards and considering the signals of the SCADA system from a statistical point of view. The creation of the fuzzy systems implies the definition of a set of fuzzy rules.

In this case, the variables considered correspond to the distance of the value measured by the SCADA to the simple moving average. The more distance to the average, the more probability of being an abnormal measure, and therefore, the more probability to generate an alarm.

Three different outcomes of the fuzzy system have been considered. Firstly, the values are in range of normal behavior and no actions are required. Secondly, orange alarms where the probability of alarm reaches exceeds a defined threshold but it is not a critical point. Finally, red alarm when the probability is unacceptable and the system need an urgent action.

The results of this methodology can become a statistical support for the generation of alarms. The methodology can be also used as a complementary information for evaluate the priority of each alarm.

Acknowledgements. The work reported herewith has been financially supported by the Spanish Ministerio de Economáa y Competitivid ad, under Research GrantsDPI2015-67264-P and RTC-2016-5694-3.

References

1. Abreu GD, Ribeiro FJ (2003) On-line control of a flexible beam using adaptive fuzzy controller and piezoelectric actuators. Sba Controle & Automação Sociedade Brasileira De Automatica 14(4):319–338
2. Acciona (2017) Wind power evolved aw1500 technical specifications. http://www.acciona-windpower.es/media/1390114/11052015-aw11051500_inusa_abril-11052012.pdf
3. Benmessaoud T, Mohammedi K, Smaili Y (2013) Influence of maintenance on the performance of a wind farm. Przeglad Elektrotechniczny 89(3):174–178
4. Camacho E, Requena V et al (2014) Demonstration of methods and tools for the optimisation of operational reliability of large-scale industrial wind turbines. In: International conference on renewable energies offshore renew
5. Chen B, Qiu YN et al (2011) Wind turbine scada alarm pattern recognition. In: Renewable power generation, pp 1–6
6. Errichello R, Muller J (1994) Application requirements for wind turbine gearboxes. NASA STI/recon Technical report N 95
7. Garcíaárquez FP, Chacón Muñoz JM (2012) A pattern recognition and data analysis method for maintenance management. Int J Syst Sci 43(6):1014–1028
8. González-Carrato RRDLH, Márquez FPG, Dimlaye V, Ruiz-Hernández D (2014) Pattern recognition by wavelet transforms using macro fibre composites transducers. Mech Syst Sig Process 48(1–2):339–350
9. González-Carrato RRDLH, Márquez FPG, Dimlaye V (2015) Maintenance management of wind turbines structures via MFCS and wavelet transforms. Renew Sustain Energy Rev 48:472–482
10. Khanna V, Cheema RK et al (2013) Fire detection mechanism using fuzzy logic. Int J Comput Appl 65(12):82–97
11. Marquez FPG (2006) An approach to remote condition monitoring systems management. In: The institution of engineering and technology international conference on railway condition monitoring, pp 156–160

12. Márquez FPG, García-Pardo IP (2010) Principal component analysis applied to filtered signals for maintenance management. Qual Reliab Eng 26(6):523–527
13. Márquez FPG, Pedregal DJ, Roberts C (2013) New methods for the condition monitoring of level crossings. Int J Syst Sci 46(5):878–884
14. Marugan AP, Marquez FPG (2016) A novel approach to diagnostic and prognostic evaluations applied to railways: a real case study. Proc Inst Mech Eng Part F J Rail Rapid Transit 230:271–284
15. Marugán AP, Márquez FPG, Pérez JMP (2016) Optimal maintenance management of offshore wind farms. Energies 9(1):46
16. MathWorks (2007) Fuzzy logic toolbox. https://es.mathworks.com/products/fuzzy-logic.html
17. Muñoz CG, Márquez FG (2016) A new fault location approach for acoustic emission techniques in wind turbines. Energies 9(1):40
18. Muñoz CQG, Márquez FPG, Tomás JMS (2016) Ice detection using thermal infrared radiometry on wind turbine blades. Measurement 93:157–163
19. Muñoz JMC, Márquez FPG, Papaelias M (2013) Railroad inspection based on ACFM employing a non-uniform b-spline approach. Mech Syst Sig Process 40(2):605–617
20. Qiu Y, Feng Y et al (2012) Wind turbine scada alarm analysis for improving reliability. Wind Energy 15(8):951–966
21. SKF (2013) SKF spherical roller bearings, setting the standard for performance and reliability. SKF. http://www.skf.com/binary/143-148465/146100_EN.pdf
22. Zhang P (2010) Advanced Industrial Control Technology

On Causal Analysis of Accident and Design of Risk-Proof Procedure for Nuclear Materials Operation: The Case of JCO Accident

Sachiko Oshima(✉)

Faculty of Science and Engineering, Waseda University, Tokyo, Japan
oshima.sachiko@aoni.waseda.jp

Abstract. This paper aims to examine a risk-proof procedure for nuclear materials operation through the causal analysis of JCO accident in 1999. Nuclear chain reactions are, by now, commonly used in the nuclear reactors, and thus it seems that there is no basic problem in fission processes from the scientific point of view. However, the criticality accident that occurred in JCO suggests that one should carefully examine this accident from the nuclear physics point of view. Indeed the nuclear chain reactions should have taken place in the small area of sedimentation tank. In fact, when people carry the uranium nitrate solution into the tank, then this solution with uranium should get into the critical state at given uranium nitrate solution. The root cause of the accident should not be very simple from the nuclear physics point of view, and it should be quite important to examine why the uranium nitrate solution could have become critical. Based on the analysis result, this paper also proposes the conditions for the operation without criticality accident which leads to serious damage of workers' life.

Keywords: Risk management · Operation system design · Nuclear fission · Criticality accidents

1 Introduction

In this paper the contents and/or expressions of the previous paper published by the first author [8] is reused for readers' convenience and understanding. The criticality accident that occurred in JCO company in 1999 (called Tokai-mura JCO-rinkai-jiko) in Japanses must be most serious, and it should not be very easy to understand why the nuclear chain reactions could proceed in a small area of sedimentation tank for a finite period of time. In this sense, it should be quite important to carry out the careful examination of criticality accidents from the nuclear physics point of view. It should be, of course, difficult to claim that the JCO accident can be a target of the scientific study as it is impossible to make the experimental study of the JCO type accidents. However, we believe

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_62

that the basic mechanism of the criticality accident should be clarified why it could naturally occur in the small area.

Here, we explain briefly the theoretical examination of the criticality accident, which are carried out in terms of the multiple scattering theory, and we show why the nuclear chain reactions can proceed in the small area [8]. In particular, the nuclear fission reactions (nucleon-nucleon collision together with nuclear fission) is traced each by each, and the microscopic processes why and how the criticality accident occurred is clarified. As a result, this paper provides some specific reasons why the chain reactions can proceed, and this can be done by making use of the mean free path which is the result of the nuclear multiple scattering theory.

When the causal event of this criticality accident is identified, another question as to why the criticality stop must be analyzed. This study tries to find an answer for this question, though not necessarily sufficient. This mechanism of stopping criticality may be related to the quick settle of the uranium compound. The calculation of this paper tells a possible dangerous situation which was thought to be due to the eighth batch, if it were carried into the sedimentation tank. The estimated energy release after the virtual eighth batch should become the same order of magnitude as the Chernobyl nuclear accident.

Finally, a procedure is designed to prevent in advance criticality accidents from the standpoint of three manufacturing resources that is Human, Machine and Material [11], grounded in the above physical theoretical discussion.

2 Research Background

2.1 Nuclear Chain Reactions

Nuclear fission reaction by incident neutrons can be written as Bohr et al. [1].

$$n + {}^{235}U \rightarrow A_1 + A_2 + (2 \sim 3)n, \tag{1}$$

where A_1, A_2 are new nuclei which are produced in the reactions. In this reaction, there are two important points. The first one is concerned with two or three neutrons which are produced in the reactions. The second point is that the probability of this nuclear reactions is strongly based on the incident neutron energy, and the biggest cross section is for the incident neutron with almost zero energy (thermal energy).

The chain reactions indicate that the produced neutrons should be absorbed by another ${}^{235}U$ such that the nuclear fission can proceed further on. In addition, if the chain reactions continue to proceed without the aid of other external neutron sources, then this situation is called a criticality stage. In reactors, this criticality must be kept by controlling the number of neutrons involved in the chain reactions.

In normal reactors, a few percent enriched uranium (${}^{235}U$ exists a few percent of total amount) should be commonly used, but in this JCO accident, 18.8% enriched uranium were used, and this high enrichment should be one of the strong reasons why the nuclear reactions run wild.

2.2 JCO Criticality Accident

This criticality accident occurred when workers in JCO were carrying the uranium nitrate solution (18.8% enriched uranium) into sedimentation tank [5,9,10]. They executed the following procedure. First, they made the uranium nitrate solution which was composed of 2.4 kg U_3O_8 with the nitric acid of 1.7ℓ in the stainless vessel. In addition, they add water to the uranium nitrate solution until the total volume became 6.5ℓ. Then, they carries the 6.5ℓ solution into the sedimentation tank, and this working procedure was called one batch. The criticality accident should have occurred in the middle of the seventh batch since the workers noticed blue lights that should be due to the Cherenkov radiation. In fact, two of the workers suffered from the neutron radiation.

A question should arise as to how the nuclear chain reactions could proceed within the small sedimentation tank (45 cm diameter, 60 cm high). There are, of course, some analyses of this criticality accident [10]. However, these studies are mainly carried out by the computer simulation such that the total energy emitted via radiations can be reproduced in some way or the other. These investigations are, of course, very important in order to understand the accident cause. However, it is also important to carry out the study of the criticality accident from the nuclear physics point of view.

3 Causal Analysis of JCO Case

3.1 Why Criticality?

Now, a question is as to why the criticality is realized in the small area of the sedimentation tank with 50ℓ of the uranium nitrate solution. That is, why nuclear chain reactions continue to occur in this small area. Here this section clarifies the basic mechanism of the criticality accident.

(1) Neutron source

The nuclear chain reactions should require thermal neutrons to start for the initial fission reactions. Since neutrons should decay within 15 min, they do not exist as a natural source. Neutrons should be produced in some way or the other. Here in this accident, the neutron source should be the decay of ${}^{238}U$ spontaneous fissions. The life time of ${}^{238}U$ is about 4.5 billion years and, in addition, the rate of the spontaneous fission to the total width is around 5.45×10^{-7}. Therefore, 1 gram weight of ${}^{238}U$ make the spontaneous fission of 0.01 times per second. Since one batch contains 1.6 kg of ${}^{238}U$, we should find about 20 neutrons per second in the one batch solution.

(2) Mean free path of $n - {}^{235}U$ fission (fast neutrons)

The probability of nuclear fission of ${}^{235}U$ induced by neutrons should be evaluated in terms of mean free path of λ inside the uranium nitrate solution. This mean free path of nuclear reactions can be obtained from the multiple scattering theory as

$$\lambda_f = \frac{1}{\rho \sigma_f}. \tag{2}$$

This derivation of the mean free path (2) is based on the Glauber theory [4], and this theoretical frame work is well examined in atomic and nuclear reactions [2,3]. Here, ρ denotes the number density of ^{235}U in solution and σ_{f} corresponds to the nuclear fission cross section of ^{235}U induced by neutrons. In fact, the number density of ^{235}U in one batch solution is $\rho \approx 1.5 \times 10^{20}$ numbers/cm^3 which is a constant. On the other hand, the nuclear fission cross section σ_{f} of ^{235}U induced by neutrons crucially depends on the incident energy of neutrons. The incident energy dependence of the observed cross sections σ_{f} can be written as [7].

$$\sigma_f \approx \begin{cases} 585\mathrm{b}: & E_n \approx 0.025\mathrm{eV} \\ 1\mathrm{b}: & E_n \approx 1\mathrm{MeV}, \end{cases} \tag{3}$$

where $1\mathrm{b} = 10^{-24}$ cm^2.

(3) Mean free path of prompt neutrons in nuclear fission

In fission process, the average energy of prompt neutrons is around 1 MeV, and therefore the average mean free path of the prompt neutrons after fissions becomes

$$\lambda_f = \frac{1}{\rho\sigma_f} \approx 67 \text{ m}. \tag{4}$$

This is quite long in comparison with the scale of the tank, and therefore this prompt neutrons by themselves cannot induce subsequent fissions in corresponding solution in the tank. In this respect, we ask a question as to why the criticality should take place within the small sedimentation tank.

3.2 Collision Between Neutrons and Water Molecule

In reality, the prompt neutrons may collide with protons in water molecule, and they should lose their energy by nucleon-nucleon collisions. Since the nuclear fission cross sections become largest for the thermal neutrons, the fission processes should start in case the prompt neutrons lose most of their energy inside the uranium nitrate solution.

(1) Energy Loss after the Collision of Prompt Neutrons with Protons in Water

When the prompt neutron scatters with protons in water, this neutron should lose a half of its energy. This can be easily understood in the following way. First, we denote the incident momentum and energy of the neutron by p, $E_n(= p^2/2M)$, and the final momentum and energy by k, $\mathrm{E_n}'$ with $E_n{}'(= k^2/2M)$. In this case, we find an equation from the conservation law of momentum and energy as

$$\frac{P^2}{2M} = \frac{k^2}{2M} + \frac{(p-k)^2}{2M}, \tag{5}$$

which can be solved and its solution becomes

$$k = p\cos\theta. \tag{6}$$

Since the observed scattering cross section does not depend on the scattering angles, we can make an average over the angles, and we obtain the average energy after the scattering

$$E_n{}' = \frac{1}{\pi}\int_0^{\pi} \frac{k^2}{2M} d\theta = \frac{1}{\pi}\int_0^{\pi} \frac{p^2}{2M}\cos^2\theta d\theta = \frac{1}{2}E_n. \tag{7}$$

This means that a neutron should lose a half of its energy in each scattering process.

(2) The Mean Free Path of Neutrons inside Water

Now we calculate the mean free path of neutrons after the scattering with protons in one batch solution. The number density of protons in one batch solution is $\rho_p \approx 4.9 \times 10^{22}$ numbers/cm^3. The neutron-proton cross section at low energy is observed as $\sigma_{np} \approx 20b$ [6] and thus the mean free path of neutron in one batch solution becomes

$$\lambda_p = \frac{1}{\rho_p \sigma_{np}} \approx 1\text{cm}. \tag{8}$$

Therefore, a prompt neutron with 1 MeV energy should have its energy after it travels around 25 cm,

$$E_n{}' = 1\text{MeV} \times \left(\frac{1}{2}\right)^{25} \approx 0.03\text{eV}. \tag{9}$$

This neutron does not have to travel linearly, but in any case, it should become a thermal neutron.

(3) Mean Free Path of Thermal Neutron in the $n - {}^{235}U$ Fission Process

We can easily calculate the mean free path of the thermal neutron before the nuclear fission in one batch solution. Since $\sigma_f = 585$b, we find

$$\lambda_f = \frac{1}{\rho \sigma_f} \approx 11 \text{ cm}. \tag{10}$$

From these considerations, we see that prompt neutrons with 1 MeV should travel around 25 cm, and then they become thermal neutrons. Further, after they travel 11 cm, they can induce nuclear fissions. Thus, if one carries 50ℓ of the uranium nitrate solution into the sedimentation tank with 45 cm diameter and 25 cm height, then nuclear chain reactions may well start quickly and proceed further on.

(4) Reaction Time of Neutrons

Now we see that when prompt neutrons travel 36 cm, then they can induce nuclear fissions. Therefore, we should estimate the duration time that is necessary to travel this 36 cm. Since the nuclear reaction time must be smaller than 10^{-15} second, we can ignore this time duration. Since the prompt neutron with 1 MeV should spend $\tau^0 \approx 7.6 \times 10^{-10}$ second to proceed 1 cm, its energy becomes a half of the previous energy after 1 cm walk. Therefore, the time to proceed the next 1 cm becomes larger by a factor of $\sqrt{2}$. In this way, if the prompt neutron proceed 25 cm, then the total time to spend must be

$$T_0 = \left(1 + \sqrt{2} + \cdots + 2^{\frac{25}{2}}\right) \tau_o \approx 15\,\mu\text{s}. \tag{11}$$

After that, this neutron becomes thermal, and it should proceed 11 cm before the nuclear fission. Since the thermal neutron may have the energy of 0.03 MeV, it should take $\tau_{th} \approx 46\,\mu s$. Thus, the total time that is necessary for the prompt neutron to induce a fission reaction should be $T_{\text{total}} \approx 61\,\mu s$.

3.3 Total Energy of Fission with Criticality

Here, we should estimate the total amount of energy which is released from this accident. This evaluation must be very difficult, but we want to calculate it in an approximate way and obtain an order of magnitude of the total energy. First, the number of neutrons which is required for the criticality reactions should be taken as $n_r = 1.001$, which is assumed to be consistent with the total energy released as calculated from the computer simulation. In nuclear reactors, one should make use of all the possible techniques to keep the number as $n_r = 1$. In addition, we assume that the number of nuclear fissions should be $N = 40000$. This number is chosen so that the total nuclear energy release should be consistent with the computer simulation which can reproduce all the observed radiation energies. In this case, the total reaction time of fission becomes $T_f \approx 2.4s$, and the total number of fissions becomes

$$N_{\text{total}} = 1.001^{40000} \approx 2.3 \times 10^{17}. \tag{12}$$

Further, we evaluate the neutron number at the beginning, and this neutron should come from the spontaneous fission of ^{238}U. The number of neutrons in one batch solution must be around 20, and a half of them are assumed to contribute followed reaction. The energy release from the nuclear fission must be around 200 MeV in each reaction, and therefore the total energy becomes

$$E_{\text{total}} \approx 4.6 \times 10^{26}\text{eV}, \tag{13}$$

which is just similar to the result of the computer simulation.

3.4 Why Does the Criticality Stop?

It is true that the criticality accident produced a huge amount of energy by the nuclear chain reactions, and the accident is indeed quite serious. In this sense, we here clarify as to how the chain reactions started and continued by reaching the critical stage. However, we face to the more serious problem at this point. That is, *why the criticality accident could stop?* We should understand any reason why the criticality could stop, namely there were only one burst and not any more burst, but why?

(1) Nuclear Fission in the Seventh Batch

Here, we try to answer for this question, though it should be extremely difficult. In order to find a possible mechanism for the stopping of the criticality, we assume that the uranium compound should settle faster than any other compounds in the solution. Further, we assume that uranium should be settled within 20% height from the bottom of the sedimentation tank.

In this case, after the sixth batch, the uranium should be settled up to the 4.9 cm from the bottom. Thus, water should be found for 19.7 cm long in the sedimentation tank. By taking into account this fact, we can calculate the total energy release by nuclear fission as

$$E_{\text{total}} \approx 4.6 \times 10^{26}\text{eV} \approx 7.4 \times 10^{7}\,\text{J}. \tag{14}$$

The duration time of this nuclear reactions can be estimated and should be around $T_f \approx 2.4\,\text{s}$, which should correspond to the time that the uranium compound is coming down to the bottom.

(2) Nuclear Fission in the Sixth Batch

The same calculation can be carried out for the sixth batch case. In this case, we see that the total energy must be 1000 times smaller than that of the seventh batch case. This is not very large, but at the sixth batch, the nuclear chain reactions already started, and indeed there were a small burst.

From this calculation, we now understand the reason why the criticality stopped. In case the uranium were settled at the bottom of the tank, then the nuclear chain reaction cannot proceed further since the prompt neutrons cannot lose their energy because of the lack of water.

(3) Nuclear Fission in the Eighth Batch

From now on, we only present a possible scenario of nuclear accident, if the eighth batch were carried into the tank. In this case, the number of uranium involved in the nuclear fission must be proportional to the height of water, and thus it should be 22.9/19.7 more than the seventh batch. Thus, the number becomes

$$N = 40000 \times \frac{22.9}{19.7} \approx 46500. \tag{15}$$

This means that the number of nuclear fissions should be also increased and the total number becomes

$$N_{\text{total}} = 1.001^{46500} \approx 1.5 \times 10^{20}. \tag{16}$$

Therefore, the total energy becomes

$$E_{\text{total}} = 3 \times 10^{29}\text{eV} \approx 4.8 \times 10^{10}\,\text{J}. \tag{17}$$

This energy 4.8×10^{10} J corresponds to 11 ton of TNT powder which is quite a serious explosion. The accident of Chernobyl nuclear power plant is believed to correspond to around 100 ton of TNT powder, and therefore, if the eighth batch were thrown away, then the accident would have been more than serious.

4 Design of Risk-Proof Procedure

Table 1 illustrates a procedure, including control factors of the manufacturing system of uranium compound and objectives, designed to prevent criticality accidents in advance from the standpoint of three manufacturing resources that is Human, Machine and Material, grounded in the above theoretical discussion. Proposed procedure mainly aims to reduce the opportunity of accelerated spontaneous occurrence of fission.

For the viewpoint of human, a control factor is the number of times workers carry uranium nitrate into sedimentation tank. Whenever this operation is activated, the neutrons settled in the tank should not contact with water molecule contained in uranium nitrate solution. The opportunity between the two materials must be minimized if possible.

For the viewpoint of a machine, a control factor is the space of sedimentation tank. The shape of the tank id better to be shallow with thinly fill of water because neutrons will be severely scattered in this case. Owing to it, the moving distance of neutrons in the water of tank can be minimized.

For the viewpoint of a material, two control factors are identified which can regarding to two components of uranium nitrate solution i.e. neutrons and water molecule. The former is the core of fission. The more the number of neutrons are, the more the opportunity of fission of neurons increases. It is necessary to control suitable number density of ^{235}U. The latter has the role to reduce

Table 1. Risk-proof procedure

Manufacturing resources	Control factors	Objectives
Human	The number of times workers carry into sedimentation tank	Minimization of the opportunity of neutrons to make contact with water molecule
Machine	Space of sedimentation tank	Minimization of moving distance of neutrons into the water in sedimentation tank
Material	The number density of ^{235}U	Reduction of the probability of induced fission by neutrons
	The amount of water	Prevention of over- slowdown material

energy of neutrons which lead to accelerated fission. Over-input of water may cause diversification of neutron occurrence and losing of their energy. Suitable amount of water should be fixed to prohibit any accident.

The above conditions were derived from consideration on operational resources. In the future, the comprehensive design of effective countermeasures are needed. The accident focused in this paper is old case, however, the proposal approach that intends to develop the procedure based on the principle of physical phenomenon will be useful for experienced IEer.

5 Concluding Remarks

The purpose of this study aims to examine a risk-proof procedure for nuclear materials operation through the causal analysis of JCO accident in 1999. In the analysis, the basic mechanism of the accident is discussed in terms of the nuclear multiple scattering theory. As the result, some specific reasons why the chain reactions can proceed and why the criticality could stop. The control factors of operation resources are extracted referred to physical phenomena regarding to fissions.

This paper aims to assist managers and academic researchers in the area of atomic power generation. As atomic power plants rely on many human operations, mangers have to cope with operational effectiveness and safety. Therefore, they have to concern to provide systematic procedure based on the proposed methodologies.

Especially, this paper proposed control factors to prevent accident. Subsequently, because of its early stage, we hope related studies to this approach is advanced by researchers interested in various areas.

Acknowledgements. We, the authors, sincerely acknowledge Dr T. Fujita, the supervisor of Doctor of Science Degrea of the 1st author, for his kind suggestions and support on preparation of this paper.

References

1. Bohr A, Mottelson BR (1998) Nuclear Deformations. Nuclear Structure
2. Fujita T, Hüfner J (1979) Inclusive hadron-nucleus scattering at high energy. Phys Lett B 87(4):327–330
3. Fujita T, Hüfner J (1980) Momentum distributions after fragmentation in nucleus-nucleus collisions at high energy. Nucl Phys A 343(3):493–510
4. Glauber RJ (1959) Lectures in Theoretical Physics. Interscience
5. JCO accident investigation committee (2005) Criticality accident in JCO; its solution, fact, cause and study (jco rinkaijika; sonozenbou no kai jijitsu youin taiou). In: Iinkai JJC, Gakkai NG (eds) Atomic Energy Society of Japan. Tokai University Press, Kanauawa
6. Nuclear Data Center (2016). http://wwwndc.jaea.go.jp/j40fig/jpeg/h001_f1.jpg
7. Okajima S, Kugo T, Mori T (2012) Reactor Physics (Genshiro Butsurigaku). Ohmsha, Tokyo

8. Oshima S, Hisada M et al (2016) Reexamination of criticality accident in JCO. J Mod Phys 07(12):1500–1505
9. Tateno J, Noguchi K, Aoyagi N (2000) Criticality Accident in Tokaimura; Complete Analysis (Tetteikaimei Tokaimura Rinkaijiko). Shin-nihon Shuppan-sha, Tokyo
10. Tonoike K, Nakamura T et al (2003) Power profile evaluation of the jco precipitation vessel based on the record of the gamma-ray monitor. Nucl Technol 143(3):364–372
11. Yamauchi K, Murata K, Katayama H (2010) Reinforcement of performance competitiveness by concurrent strategy: a case of chemical product manufacturing line. In: Proceedings of the 23rd International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM 2010), pp 913–920

Multi-stage Logistics Inventory for Automobile Manufacturing by Random Key-Based GA

Hisaki Inoue[1(✉)], Jung Bok Jo[2], and Mitsuo Gen[3]

[1] Miyazaki Sangyo Keiei University, Miyazaki 880-0931, Japan
inoue-lab@sankei.info
[2] Division of Computer Engineering, Dongseo University, Busan, Korea
[3] Fuzzy Logic Systems Institute, Tokyo University of Science, Tokyo, Japan

Abstract. When evaluating a logistics system, automobile companies commonly search for the minimum transportation cost, which is significantly influenced by inventory problems. These inventory problems are extensive and varied. In many actual logistics systems, there are the three-stage network models take inventory values into consideration. Safety inventories are kept in distribution centers (DCs). In this study, we adapted a model to set-up a number of plants and DCs. We then performed numerical experiments by using demand data that we created on the basis of data disclosed by an automobile company. In this study, we propose a random key-based genetic algorithm (rk-GA) with the distributed environment scheme, we compared it with random key and spanning tree-based GAs, and report the advantages of the proposed method, random key-based genetic algorithm with distributed environment scheme (des-rkGA).

Keywords: Automobile manufacturing · Multi-stage logistics · Inventory control · Random key-based genetic algorithm

1 Introduction

Many difficult inventory problems involve production control and asset management, but those that involve logistics systems are important and have garnered strong interest in recent years. Particular focus has been paid to inventory problems. Doboshas presented many papers on inventory problems. In 2001, he presented the problem of reverse logistics, adjusting the relationship of holding, production, and disposal costs [4]. In 2005, he investigated production inventory adjustment [5]. In 2007, he presented a paper on the total production cost of two companies for adjustment [6]. Minner et al. [18] have also presented papers on inventory problems. In 2003, they evaluated reverse logistics, where the inventory has several supply methods. In 2005, they presented a paper on the problem of adjusting shipping, replenishment, and lost sales opportunities for two inventories [19]. In 2008, Thangam et al. [28] presented a paper on how to determine

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_63

replenishment for Poisson demand. In 2003, Mahadevan et al. [16] treated a facility as an inventory problem where the returned products are remanufactured. In 2004, Miranda et al. [20] analyzed the inventory decision problem using the Lagrangian relaxation method and subgradient methods; their ordering point method was based on the economic order quantity. In 2009, Rieksts et al. [24] analyzed the inventory problem with ordering intervals using power- of- two policies.

There are many other inventory problems, including the reduction of the total safety inventory quantity, or on-hand inventory [8], and the calculation of the inventory value at each step, or echelon inventory [12].

In this study, we propose a new model where inventories are managed at distribution centers (DCs), taking actual conditions into account. Holding costs only occur in DCs, but additional inventory costs are incurred for the product value because a product near dealers has more value. For inventory costs other than the holding cost, we can get an interest charge if the product is exchanged with cash. Another factor is the lost product value due to age depreciation. We calculated the annual supply and demand value, and created demand data that were based on the Poisson demand of time-series fluctuations. Logistics models such as these are known as the NP-hard problem [3]. Soft computing methods such as simulated annealing, neural networks, and genetic algorithms (GA) are well-suited to solve this problem [1,17,26,29].

In this study, we adopted random key-based genetic algorithm (rk-GA) with distributed environment scheme (des-rkGA) as the proposed method, which is an improved version of rk-GA; we compared the proposed method with rk-GA and spanning tree-based GA (st-GA) to confirm its suitability [9].

We propose a model that addresses many of the different inventory problems studied earlier; we demonstrate des-rkGA algorithm to solve the multi-logistics inventory problem, and present the computational results with the effectiveness by the proposed algorithm.

2 Inventory Problem in Logistics System

In this section, we detail the contents for the inventory problem treated in this study. Inventory costs result from holding costs and other factors. For the model used in this study, the product is held in a DC for general managing. All products have needed costs as their product value, and the production control cost is needed to calculate the safety inventory in a DC. The safety inventory quantity is decided by the service level but is affected by the production control and holding costs. The safety inventory cost is based on the service level and has a trade-off relation with lost costs based on lost sales opportunities.

Figure 1 shows many different kinds of inventories; they are difficult to categorize clearly, but it is important for inventory management to categorize and manage them. The inventory cost without the holding cost includes all inventory, such as the pipeline, production process, lot size and DC inventories. The inventory costs determine the product value, inventory value, inventory time, and inventory holding ratio.

Therefore, the inventory cost without holding costs and the product value, inventory value, inventory time, and inventory holding ratio have a trade-off relation. In this section, we describe the inventory holding ratio, safety inventory, production adjustment, pipeline inventory, and inventory on a production process, lot size inventory, and DC inventory as well as a summary of the inventory problems evaluated in this study. Although there are many kinds of inventories, we adopted the idea of a value chain for the inventories; we used the pipeline, lot size, production process, and DC inventories.

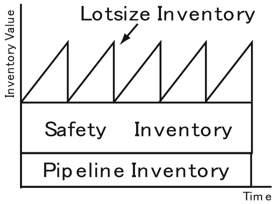


Fig. 1. Inventory category

The general outline of the logistics model used for the inventory in this study is shown in Fig. 2.

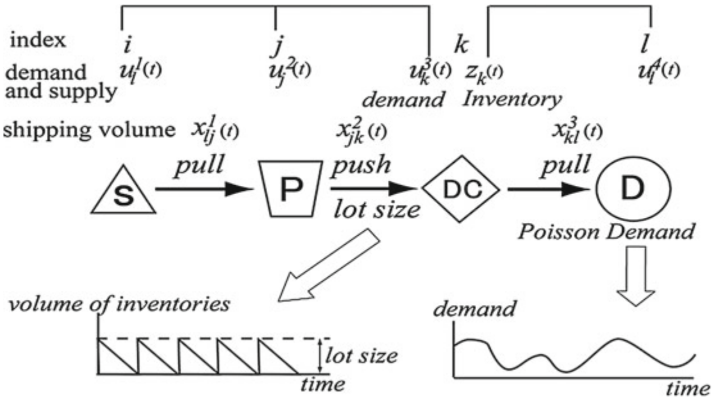


Fig. 2. General outline of logistics model with inventory

Below, we define the indices, constants, parameters, decision variables, and units used in this paper. Units of measurement are in square brackets.

Indices:

$i = 1, 2, \cdots, I$: index of suppliers;
$j = 1, 2, \cdots, J$: index of plants;
$k = 1, 2, \cdots, K$: index of DCs;
$l = 1, 2, \cdots, L$: index of customers;
$t = 1, 2, \cdots, T$: index of cycles in the logistics system. T_{cycle} intervals are described later.

Parameters:

A : number of unit parts to constitute;
c^1_{ij} : shipping cost of unit parts or material from supplier i to plant j [yen];
c^2_{jk} : shipping cost of unit production from plant j to DC k [yen];
c^3_{kl} : shipping cost of unit production from DC k to customer l [yen];
g^1_{ij} : shipping time of unit parts or material from supplier i to plant j [h];
g^2_{jk} : shipping time of unit production from plant j to DC k [h];
g^3_{kl} : shipping time of unit production from DC k to customer l [h];
T_{plant} : producing time of plants [h];
α : safety inventory coefficient;
σ : standard deviation of demand;
T_{cycle} : cycle time in the logistics system. It shows how much time it takes for a load to be moved once [h];
h : inventory holding ratio (2.1 value chain.);
H : holding cost of unit production in a DC [yen];
e_k : inventory cost or DC k [yen];
M_{cost} : material cost [yen];
r^1_j : fixed cost for operating plant j [yen];
r^2_k : fixed cost for operating DC k [yen];
U^1_i : upper limit of supply for parts and materials in supplier i;
U^2_j : upper limit of supply for production in plant j;
U^3_j : upper limit of supply for production in DC k;
p^{const}_j : unit cost production at steady state [yen];
p^{exceed}_j : unit cost production at excess state, [yen];
p^{shortage}_j : unit cost production at shortage state [yen];
I_{total} : total inventory cost without holding cost;
N_{delay} : number of delays for plant production;
W_1 : number of plants that can be operated;
W_2 : number of DCs that can be operated;
R_{average} : production quantity for one period of cycle time as calculated from the annual average production quantity;
S^1_i : supplier production ratio; ratio of supplier i to gross supplier product capability;
S^2_j : plant production ratio; ratio of plant j to gross plant product capability;
z^{level}_k : base inventory level in DC k.

Decision Variables:
$u_i^1(t)$: supply amount of parts and materials in supplier i at cycle t;
$u_j^2(t)$: supply amont of production in plant j at cycle t;
$u_k^3(t)$: supply amont of production in DC k at cycle t;
$u_l^4(t)$: demand amount of production in customer l;
$z_k(t)$: inventory volume for DC k at cycle t [yen];
$S_k^3(t)$: DC demand ratio; demand ratio of DC k to the gross DC demand quantity;
$R_j^2(t)$: shipping quantity in plant j;
$R_k^3(t)$: receive cargo quantity in DC k;
$B_l(t)$: back order quantity in customer l;
$Dl(t)$: order quantity at this period in customer l;
$D_l^{DC}(t)$: request quantity to DC at this period in customer l;
$\Delta z_k(t)$: difference of inventory quantity and base inventory level;
$\Delta z_{\max}(t)$: maximum of difference of inventory quantity and base inventory level;
$z_k^{\mathrm{req}}(t)$: request quantity in DC k;
p_{val}^1 : product value when delivery is completed to plants [yen];
p_{val}^2 : product value when delivery is completed to DCs [yen];
p_{val}^3 : product value when delivery is completed to customers [yen];
U_j^{exceed} : threshold for determination when exceeding production;
U_j^{shortage} : threshold for determination when reducing production;

Decision Variables:
$p_j(uj^2)$: producing cost of unit production in plant j, [yen]
$x_{ij}^1(t)$: amount supplied of unit parts or material from supplier i to plant j at cycle t;
$x_{jk}^2(t)$: amount supplied of production from plant j to DC k at cycle t;
$x_{kl}^3(t)$: amount supplied of production from DC k to customer l at cycle t;
$p_j^1(t)$: operating flag for plant j at cycle t ($= 1$ when plant j is used, $= 0$ otherwise);
$p_k^2(t)$: operating flag for DC k at cycle t ($= 1$ when DC k is used, $= 0$ otherwise).

2.1 Value Chain

When considering the product value and inventory holding ratio in addition to the holding cost, the inventory cost becomes high near the customer (dealer), as shown in Fig. 3 [23]. The inventory holding ratio is determined by the constant number by interest rate when the product is cashed, the decrease in product value, etc. In this study, we treated the inventory holding ratio as uniform because we were dealing with engineered products such as automobiles that do not experience degradation. We adopted the value chain concept for all products.

The value chain can be defined as:

I_{cost} : Inventory Cost;
P_{value} : Product Value;
I_{volume} : Inventory Amount.

$$I_{\text{cost}} = h \times P_{\text{value}} \times I_{\text{volume}} \tag{1}$$

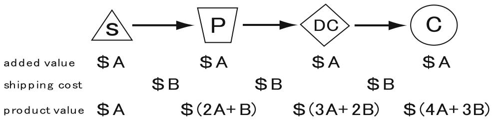


Fig. 3. Value chain

2.2 Safety Inventory

The safety inventory is the inventory required to prevent stock out. The safety inventory and service level have a mutual trade-off relation. The formula for the safety inventory is shown below.

$$I_{\text{safety}} = \alpha\sigma\sqrt{T_{\text{cycle}} + \max(g_{jk})} \quad \forall j, k, \tag{2}$$

where I_{safety} is safety inventory.

2.3 Production Adjustment

Order fluctuations mean that factories have to adjust production. Extra charges include extra pay, etc. if production is over the steady state, and extra fixed costs for the equipments if production is under steady state. The formula for production adjustment is shown below.

$$\begin{aligned} P_j(u_j^2(\mathrm{t})) =& P_{\text{const}} + \max\{P_{\text{exceed}}(u_j^2(\mathrm{t}) - U_j^{\text{exceed}}), 0\} \\ & + \max\{P_{\text{shortage}}(U_j^{\text{shortage}} - u_j^2(\mathrm{t})), 0\}. \end{aligned} \tag{3}$$

2.4 Pipeline Inventory

The pipeline inventory denotes the inventory during shipping. There is a trade-off between the pipeline inventory and shipping cost. In this study, we constructed a model based on the idea of an inventory holding ratio. We treated each product during shipping between a supplier and plant and between a DC and customer as pipeline inventory. The formula of the pipeline inventory cost I_{pileline} is shown below.

$$I_{\text{pipeline}} = h\left\{\sum_{t=1}^{T}\left(p_{val}^{1}\sum_{i=1}^{I}\sum_{j=1}^{J} g_{ij}^{1}x_{ij}^{1}(t) + p_{val}^{2}\sum_{j=1}^{J}\sum_{k=1}^{K} g_{jk}^{2}x_{jk}^{2}(t) + p_{val}^{3}\sum_{k=1}^{K}\sum_{l=1}^{L} g_{kl}^{3}x_{kl}^{3}(t)\right)\right\}. \tag{4}$$

The cost is approximated as follows because we presumed that the product has the same value when received by any factory, DC, or dealer. The value of completed product is approximated by material and average shipping costs.

$$p_{\text{val}}^{1} = M_{\text{cost}} + \frac{\sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J} c_{ij}^{1}x_{ij}^{1}(t)}{\sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}^{1}(t)}. \tag{5}$$

The product value when shipped to DCs is approximated by its value when shipped to plants and the average shipping cost.

$$p_{\text{val}}^{2} = p_{\text{val}}^{1} + \frac{\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{jk}^{2}x_{jk}^{2}(t)}{\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{k=1}^{K} x_{jk}^{2}(t)}. \tag{6}$$

The product value when shipped to dealers is approximated by its value when shipped to DCs and average shipping cost.

$$p_{val}^{3} = p_{val}^{2} + \frac{\sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{l=1}^{L} c_{kl}^{3}x_{kl}^{3}(t)}{\sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{l=1}^{L} x_{kl}^{3}(t)}. \tag{7}$$

2.5 Inventory of a Production Process

For the inventory of a production process, we constructed a model based on the idea of an inventory holding ratio. We treated an unfinished product as inventory on a production process.

The formula is shown below. The product value on a production process changes with the progress conditions of the manufacturing processes. However, in this study, the value was approximated as the material cost and half of the production cost in the plant.

$$I_{\text{production}} = hT_{\text{plant}}\sum_{j=1}^{J} u_j^2(t)(p_{val}^{1} + P_j(u_j^2(t))/2). \tag{8}$$

2.6 Lot-Size Inventory

The lot size inventory is the inventory of the completed products during shipping. We treated completed products as lot size inventory. All products are shipped from plants to DCs in the same interval. The lot size inventory cost was calculated as half the product of the shipping values, production values in plants, and inventory holding ratio, as shown in Fig. 4; we used this formula because all products were shipped from a plant to a DC in the same time intervals. The formula is shown below.

$$I_{\text{lotsize}} = h \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{g_{jk}^2 x_{jk}^2(t)(p_{val}^1 + P_j(u_j^2(t)))}{2}. \tag{9}$$

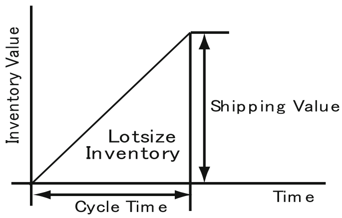


Fig. 4. Lot size inventory value

2.7 DC Inventory

We treated products in a DC as DC inventory. In this study, the DC inventory plays the central role of the inventory of the logistics system; we need extra holding costs for each car in DCs. The DC inventory cost is calculated as the product of the DC inventory amount, cycle time, and production value when shipped to DCs multiplied by the inventory holding ratio. The formula is shown below.

$$I_{DC} = hT_{\text{cycle}} p_{\text{val}}^2 \sum_{t=1}^{T} \sum_{k=1}^{K} z_k(t) + HT_{\text{cycle}} \sum_{t=1}^{T} \sum_{k=1}^{K} z_k(t). \tag{10}$$

2.8 Total Inventory Cost Without Holding Cost

As mentioned above, the formula of the total inventory cost without the holding cost used in this study consists of the pipeline, production process, lot size, and DC inventory costs; it is shown as follows.

$$I_{\text{total}}=I_{\text{pipeline}}+I_{\text{production}}+I_{\text{lotsize}}+I_{DC} \tag{11}$$

$$\begin{aligned} I_{\text{total}} =h\Bigg\{ & \sum_{t=1}^{T}(p_{\text{val}}^{1}\sum_{i=1}^{I}\sum_{j=1}^{J} g_{ij}^{1}x_{ij}^{1}(t)+p_{\text{val}}^{2}\sum_{j=1}^{J}\sum_{k=1}^{K} g_{jk}^{2}x_{jk}^{2}(t) \\ & +p_{val}^{3}\sum_{k=1}^{K}\sum_{l=1}^{L} g_{kl}^{3}x_{kl}^{3}(t))+T_{\text{plant}}\sum_{j=1}^{J}u_{j}^{2}(t)(p_{\text{val}}^{1}+P_{j}(u_{j}^{2}(t))/2) \\ & +\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{g_{jk}^{2}x_{jk}^{2}(t)(p_{val}^{1}+P_{j}(u_{j}^{2}(t)))}{2}+T_{\text{cycle}}p_{\text{val}}^{2}\sum_{t=1}^{T}\sum_{k=1}^{K}z_{k}(t)\Bigg\} \\ & +HT_{\text{cycle}}\sum_{t=1}^{T}\sum_{k=1}^{K}z_{k}(t). \end{aligned} \tag{12}$$

2.9 Demand Data

In this study, we constructed the logistics system in terms of inventories with Poisson demand [25]. To confirm the availability, we created demand data from an automobile company's public data. Many products flow in the supply chain network (SCN) by dealer's demand, and these demands are approximated with Poisson demand time fluctuations. We created new demand data using Poisson randomization. The Poisson equation is shown below.

$$p(k)=\frac{e^{-\lambda}\lambda^{k}}{k!}, \tag{13}$$

here, $p(k)$ is the demand, λ is the average order values, and k is the number of order times. Figure 5 shows the outline of the Poisson randomization. In this study, we created daily demand data each experiment by applying a Poisson random number to the annual demand.

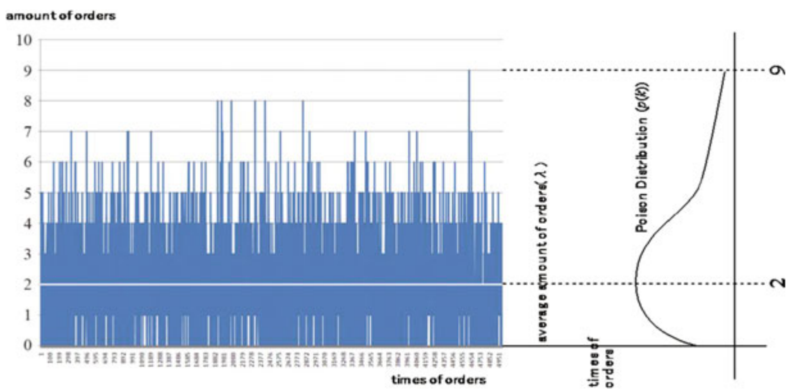


Fig. 5. Relationship between amount and times of orders

3 Mathematical Model of Multi-stage Logistics System with Inventory

3.1 Assumptions

In this study, we constructed the logistics model with the following assumptions.

A1. The transit times are known between suppliers and plants, plants and DCs, and DCs and customers.
A2. The shipping costs are known between suppliers and plants, plants and DCs, and DCs and customers.
A3. The supplies are delivered without delay to a factory in population to the planned production.
A4. In this model, suppliers provide multipurpose parts for effective optimization. The parts are examined and classified for easy assembly work. This process is carried out assuming that A package parts are used in the assembly of one car. The conditions for assembling one car require the package parts of A units. Other parts are not targeted in this model because the supply route is decided from the first time the supplier side is entrusted with delivery, few of the detailed parts have management value in the logistics system, etc.
A5. The products made at a plant are shipped to a DC by lot size.
A6. The DCs have space to accept products from plants.
A7. We consider only the inventory costs in DCs.
A8. The customer addresses are known.
A9. The existence of inventory is known in advance through an inventory check; orders for inventory that is out of stock are treated as reservations.
A10. A product has the same value when received by any factory, DC, or dealer.
A11. All products are delivered within the limits time of a T_{cycle}.

3.2 Multi-stage Logistics System

Generally, when a logistics system is seen from the functional constitutive property, it is modeled by a three-stage production network and distribution system, which is called a supply chain network. The first stage is the supplier phase, which involves parts and a supplier. The second stage is the plant phase, which consists of a production plant or outsourcing. The third stage is the DC phase and consists of a distribution center or storehouse. A sample of an actual automobile company's logistics system is shown in Fig. 6. In many cases, the actual logistics system is comprised of three stages. A three stage logistics system model is shown in Fig. 7. The logistics model used in this study had 14 suppliers, 2 factories, 4 DC, and 22 customers. The mathematical model used in this study is shown below.

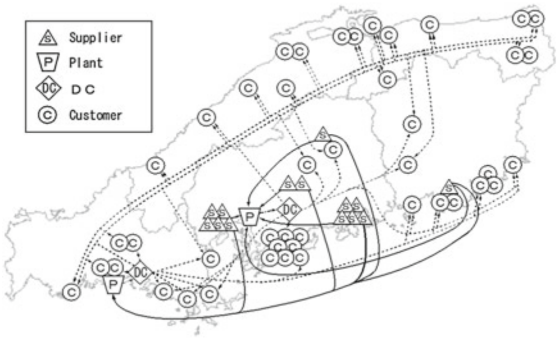


Fig. 6. Sample of actual logistics system

$$
\begin{aligned}
\min Z_1 = &\sum_{t=1}^{T}\Bigg\{A\sum_{i=1}^{I}\sum_{j=1}^{J}c_{ij}^{1}x_{ij}^{1}(t)+\sum_{j=1}^{J}\sum_{k=1}^{K}c_{jk}^{2}x_{jk}^{2}(t)\\
&+\sum_{k=1}^{K}\sum_{l=1}^{L}c_{kl}^{3}x_{kl}^{3}(t)\Bigg\}+h\Bigg\{\sum_{t=1}^{T}(p_{\text{val}}^{1}\sum_{i=1}^{I}\sum_{j=1}^{J}g_{ij}^{1}x_{ij}^{1}(t)\\
&+p_{\text{val}}^{2}\sum_{j=1}^{J}\sum_{k=1}^{K}g_{jk}^{2}x_{jk}^{2}(t)+p_{\text{val}}^{3}\sum_{k=1}^{K}\sum_{l=1}^{L}g_{kl}^{3}x_{kl}^{3}(t))\\
&+T_{\text{plant}}\sum_{j=1}^{J}u_{j}^{2}(p_{\text{val}}^{1}+P_j(u_j^2(t))/2)\\
&+\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{g_{jk}^{2}x_{jk}^{2}(t)(p_{\text{val}}^{1}+P_j(u_j^2(t)))}{2}\\
&+T_{\text{cycle}}p_{\text{val}}^{2}\sum_{t=1}^{T}\sum_{k=1}^{K}z_k(t)\Bigg\}+HT_{\text{cycle}}\sum_{t=1}^{T}\sum_{k=1}^{K}z_k(t)\\
&+\sum_{t=1}^{T}\sum_{j=1}^{J}u_j^2(t)P(u_j^2(t))+\sum_{j=1}^{J}r_j^1p_j^1(t)+\sum_{k=1}^{K}r_k^2p_k^2(t)
\end{aligned}
\tag{14}
$$

$$
\text{s. t.}\ \sum_{j=1}^{J}x_{ij}^{1}(t)\le u_i^1(t),\quad \forall i,t \tag{15}
$$

$$
\sum_{i=1}^{I}x_{ij}^{1}(t)\le u_j^2(t)p_j^1(t),\quad \forall j,t \tag{16}
$$

$$\sum_{j=1}^{J} x_{jk}^2(t) \leq u_k^3(t) p_k^2(t), \quad \forall k, t \tag{17}$$

$$\sum_{k=1}^{K} x_{kl}^3(t) \geq u_l^4(t), \quad \forall l, t \tag{18}$$

$$\sum_{j=1}^{J} x_{ij}^1(t) = \sum_{j=1}^{J} x_{jk}^2(t), \quad \forall j, t \tag{19}$$

$$\sum_{k=1}^{K} x_{jk}^2(t) = z_k(t-1) - z_k(t) + \sum_{k=1}^{K} x_{kl}^3(t), \quad \forall j, t \tag{20}$$

$$\sum_{j=1}^{J} p_j^1 \leq W_1 \tag{21}$$

$$\sum_{k=1}^{K} p_k^2 \leq W_2 \tag{22}$$

$$x_{ij}^1(t), x_{jk}^2(t), x_{kl}^3(t), z_k(t) \geq 0, \quad \forall i, j, k, l, t \tag{23}$$

$$p_j^1(t), p_k^2(t) = \{0, 1\}, \forall j, k, t \tag{24}$$

$$\begin{aligned} P_j(u_j^2) = P_j^{\text{cont}} &+ \max(P_j^{\text{exceed}}(u_j^2(t) - U_j^{\text{exceed}}), 0) \\ &+ \max(P_j^{\text{shortage}}(U_j^{\text{shortage}} - u_j^2(t)), 0) \end{aligned} \tag{25}$$

$$p_{\text{val}}^1 = M_{\text{cost}} + \frac{\sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{j=1}^{J} c_{ij}^1 x_{ij}^1(t)}{\sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{j=1}^{J} x_{ij}^1(t)} \tag{26}$$

$$p_{\text{val}}^2 = p_{\text{val}}^1 + \frac{\sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{k=1}^{K} c_{jk}^2 x_{jk}^2(t)}{\sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{k=1}^{K} x_{jk}^2(t)} \tag{27}$$

$$p_{\text{val}}^3 = p_{\text{val}}^2 + \frac{\sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{l=1}^{L} c_{kl}^3 x_{kl}^3(t)}{\sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{l=1}^{L} x_{kl}^3(t)}. \tag{28}$$

4 Advanced Genetic Algorithm

4.1 Random-Key Based Genetic Algorithm

Gen and Lin [7] surveyed genetic algorithms in Wiley Encyclopedia of Computer Science and Engineering and recently many researchers applied GA to various areas in logistics systems. Inoue and Gen [10] reported multistage logistics system with inventory considering demand by hybrid GA, Neungnatcha et al. [22] reported adaptive genetic algorithm (AGA) for solving sugarcane loading stations with multi-facility services problem, Jamrus et al. [11] reported discrete particle swarm optimization (PSO) approaches and extended priority based-HGA for solving multistage production distribution under uncertainty demands

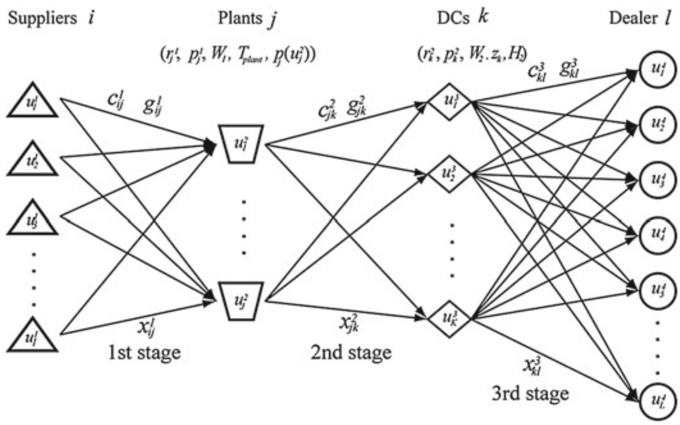


Fig. 7. 3-stages logistics model

and Lee et al. [13] reported multi-objective hybrid genetic algorithm (MoGA) to minimize the total cost and delivery tardiness in a reverse logistics.

Lin and Gen [15] proposed a random key-based genetic algorithm (rk-GA) for solving AGV (automatic guided vehicle) dispatching problem in flexible manufacturing system (FMS). Now we are going to use it for multistage logistics system with inventory. Now we define the following example of the cost matrix:

$$[e_{ij}] = \begin{bmatrix} 67 & 71 & 32 \\ 44 & 51 & 62 \\ 50 & 59 & 35 \\ 38 & 78 & 56 \\ 53 & 65 & 74 \end{bmatrix} \qquad [f_{jk}] = \begin{bmatrix} 39 & 32 & 62 & 44 \\ 45 & 56 & 53 & 59 \\ 47 & 35 & 41 & 38 \end{bmatrix}$$

$$[g_{kl}] = \begin{bmatrix} 48 & 62 & 74 & 20 & 92 & 83 & 32 \\ 51 & 26 & 89 & 17 & 35 & 59 & 86 \\ 34 & 50 & 56 & 65 & 50 & 53 & 47 \\ 71 & 44 & 23 & 38 & 11 & 62 & 29 \end{bmatrix}$$

Fig. 8. Sample of cost matrix

The algorithm created using the rk-GA technique has three logistics stages. Figure 7 shows the third stage process. Figure 8 shows a sample cost matrix. Figure 9 shows a sample rk-GA chromosome.

Gen and Cheng successfully applied rk-GA encoding to the shortest path and project scheduling problems in 2000 [2]. For transportation problems, a chromosome consists of the priorities of sources and depots, which make up a transportation tree; its length is equal to the total number of sources m and depots n, i.e., m + n. The transportation tree corresponding to a given chromosome is generated by the sequential arc between sources and depots. At each

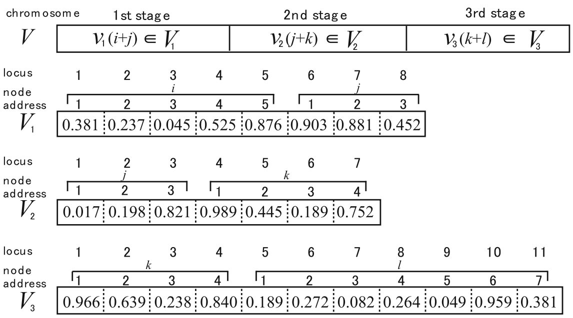


Fig. 9. Sample of rk-GA chromosome

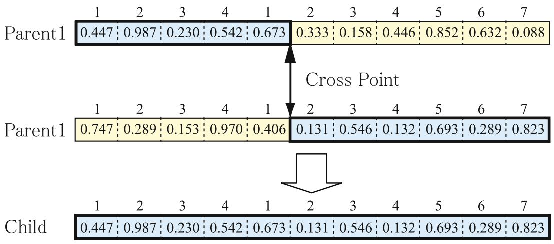


Fig. 10. Sample of one point crossover

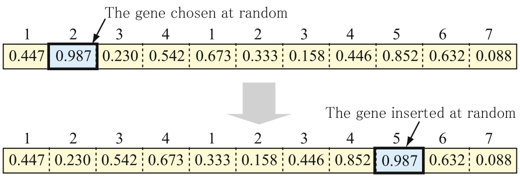


Fig. 11. Sample of insertion mutation

step, a single arc is added to a tree that selects a source (depot) with the highest priority and connects it to a depot (source) to minimize cost.

Figure 13 shows the brief decoding at each stage. Figure 14 shows the process of the m-logistics problem. In this study, we used the one-point crossover, which is the simplest method when using rk-GA. We used insertion and swap mutations. We used the roulette wheel approach, which selects the chromosome in ascending order of fitness. Examples of one-point crossover, insertion mutation, and swap mutation are shown in Figs. 10, 11 and 12, respectively.

The new generated chromosome is evaluated. It is selected in ascending order of fitness based on the number of *popSize* in the parent and newly generated chromosomes. The order of fitness then helps determine the next generation of chromosomes.

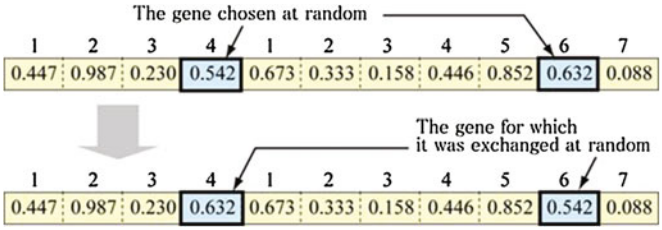


Fig. 12. Sample of swap mutation

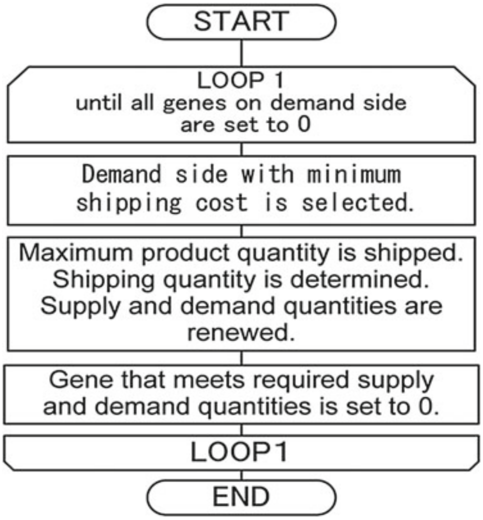


Fig. 13. Brief decoding of each stage

Inversion and displacement mutations are used in st-GA. The inversion mutation select two positions within a chromosome at random and then inverts the substring between these two positions. The displacement mutation selects a substring at random and inserts it in a random position.

4.2 Total Logistics Process

The total logistics system process can be explained as follows. Figure 15 shows the whole logistics system process in logistics cycle periods. The safety inventory is given first each cycle period. The inventory quantity is renewed last in each cycle period. The plant product shipping quantity is based on past demand, because of the production time in plants. In this study, this is called the number of delays for plant production N_{delay} where the plant production time is the number of order times. This is shown below.

```
procedure total m-logistics:
input: shipping data, GA parameter
output: minimum m-logistics cost
begin
   t ← 1;
   initialize P(t) by Random Key-based encoding routine;
   fitness eval(P) by total decoding routine;
   while (not terminating condition) do
      crossover P(t) to yield C(t) by partially matched crossover;
      mutation P(t) to yield C(t) by insertion and swap mutation;
      fitness eval(C) by total decoding routine;
      select P(t+1) from P(t) and C(t) by roulette wheel selection;
      t ← t + 1;
    end
output best m-logistics cost
end;
```

Fig. 14. m-Logistics Process

$$N_{\text{dalay}} = \frac{T_{\text{plant}}}{T_{\text{cycle}}}. \tag{29}$$

Here, we use an example for the shipping products when the customer's total demand quantity changes from 600 before delay cycle times (N_{delay}) to 1200. A pull-type demand quantity is applied to DC-customer and supplier-plant product distributes based on the demand quantity at the time. The shipping quantity is 1200. A push-type demand quantity is applied to plant-DC product distributions based on the demand quantity before the number of delay times for plant production (N_{delay}). The shipping quantity is 600.

The load of the customer's demand is shared by the total inventory quantity in DCs and total production in plants and suppliers. An example process is shown in Fig. 16. Step 1 shows a renewal of the demand quantity (u_l^4) and back order quantity (B_l). Step 2 calculates the planned order quantity ($u_j^2(t)$). Step 3 calculates the order quantity in suppliers ($u_i^1(t)$). Step 4 calculates the planned shipping quantity ($R_j^2(t)$). Step 5 calculates the quantity of cargo received in DCs ($R_k^3(t)$).

```
procedure total m-logistics in cycles period:
input: shipping data, inventory data, GA parameter
output: minimum total cost
begin
  t ← 1;
  z_k = ασ√(T_cycle + G_1);              // initialize inventory quantity
  while (not termination condition) do
    initial setting routine;      //initial setting process
    m-logistics routine;
    t ← t + 1;
   end
output best total cost;
end;
```

Fig. 15. Total m-logistics process in a cycles period

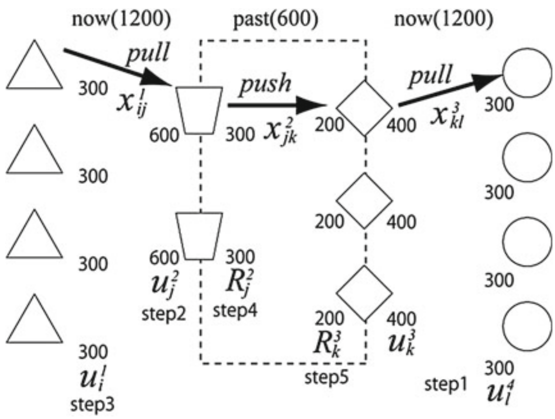


Fig. 16. Initial process

4.3 rk-GA with Distributed Environment

The immigration scheme is an independent genetic operation where each island is made up of several divided populations for the same generation. The basic concept of immigration is shown in Fig. 17. Immigration is an information exchange that is performed continuously for a group of chromosomes. The ratio to the number of immigration populations is called the immigration ratio, and the generation interval during which immigration occurs is called the immigration interval. In 2008, Lin et al. [14] solved the shortest route problem for cost and time using a priority-based GA with immigration, and reported their result.

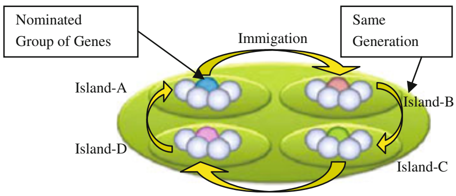


Fig. 17. Basic Concept of Immigration

In this study, we propose rk-GA with a distributed environment scheme (des-rkGA) that changes the crossover and mutation rates for each island. Miki et al. [21] proposed multi-objective genetic algorithms with a distributed environment scheme (MoGA-DES). We adopted an immigration rate of 0.1; we selected other islands randomly and changed the chromosome asynchronously. There are two methods for parallel processing [27]. The multithread method performed by one programming using time sharing. The multi-task method performs several tasks at the same time [21]. In this study, we adopted the multi-task method because it is easy to disperse processing to several PCs. There are also two processing methods for GA. The synchronous method synchronized the time for each generation. The asynchronous method does not do so. We adopted the asynchronous method because we would need a wait time for the slowest island if we adopted the synchronous method. We created nine islands with crossover probability (PC) and lower mutation probability (PM) than the center island. Two of nine islands were chosen at random with the generation timing of the center island. The direction of immigration was decided at random, and the populations immigrated at a 0.1 immigration rate. Next, 10% of each island's worst chromosomes were destroyed, and 10% of other islands' best chromosomes were adopted as immigrations. The concrete values of PC and PM used by this study are discussed in detail in the next section. We performed the experiment using parallel processing with three PCs.

5 Numerical Experiments

We performed prior experiments to determine PC and PM for the center island and obtained the solution is shown in Table 1. We adopted $P_C = 0.4$ and $P_M = 0.6$ as the center island values for st-GA and $P_C = 0.6$ and $P_M = 0.4$ as the center island values for rk-GA because they provided the best solutions.

We adopted $P_C = 0.6$ and $P_M = 0.4$ as the center island values for des-rkGA, because it is based on rk-GA. We created 9 islands with $P_C = (0.4, 0.6, 0.8)$ and $P_M = (0.2, 0.4, 0.6)$ for the experiments.

popSize denotes the population size. maxGen is the maximum generation size used as the terminating condition for the experiments. We performed experiments repeated 20 times using maxGen $=$ 5000 and popSize $= (20, 50, 100)$. However, we reported maxGen $=$ 1000 as sufficient for the evolutive process because no more improvement was detected after 1000 generations. Table 2 shows the evolutive processes for the best value by each method. The best value Z_1 was at gen $= (300, 500, 1000)$.

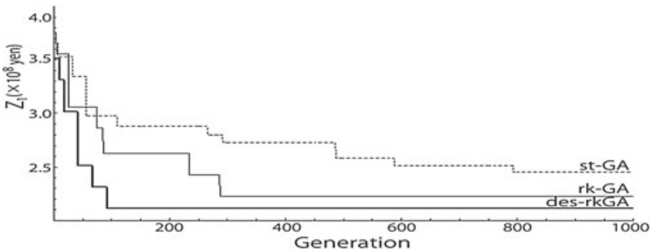


Fig. 18. Evolutive process for each method

Figure 18 shows the evolutive processes for each method when popSize was 100. The proposed Des-rkGA produced the best final result because its evolutive

Table 1. Solution at each P_C and P_M (unit: yen)

	P_M	P_C			
		0.2	0.4	0.6	0.8
st-GA	0.2	451, 461, 818	442, 315, 078	447, 088, 045	454, 269, 696
	0.4	445, 044, 804	440, 918, 667	444, 501, 564	444, 782, 298
	0.6	442, 653, 511	440, 602, 712	445, 820, 848	452, 372, 968
	0.8	445, 865, 777	443, 723, 937	451, 882, 261	454, 711, 777
rk-GA	0.2	412, 925, 351	406, 037, 962	408, 355, 139	413, 251, 110
	0.4	406, 280, 823	405, 783, 627	405, 476, 817	405, 988, 745
	0.6	408, 604, 249	406, 269, 725	405, 763, 570	413, 679, 127
	0.8	412, 120, 741	406, 777, 058	411, 116, 273	413, 073, 894

Table 2. Evolutive process for each popSize (unit: yen)

	PopSize	Gen		
		300	500	1000
st-GA	20	551, 429, 175	547, 095, 287	541, 769, 426
	50	499, 464, 718	485, 094, 520	482, 185, 269
	100	447, 629, 126	423, 064, 081	402, 412, 188
rk-GA	20	467, 164, 557	463, 804, 677	433, 653, 257
	50	452, 496, 499	434, 436, 525	394, 435, 691
	100	366, 335, 225	365, 780, 289	364, 869, 285
des-rkGA	20	418, 809, 097	396, 053, 541	389, 092, 924
	50	353, 999, 460	346, 130, 700	346, 030, 174
	100	348, 536, 438	345, 863, 047	345, 695, 925

Table 3. Evolution of each island by des-rkGA (unit: yen)

Island	Gen			
	10	20	50	70
A	590, 914, 974	505, 586, 248	442, 219, 078	407, 869, 638
B	545, 289, 874	519, 883, 420	431, 958, 888	411, 047, 134
C	596, 086, 519	530, 215, 824	413, 252, 989	399, 855, 446
D	589, 080, 440	496, 206, 544	433, 337, 814	386, 073, 858
E	569, 117, 032	535, 673, 932	419, 654, 636	412, 582, 256
F	556, 063, 137	497, 222, 107	454, 229, 946	381, 231, 373
G	582, 061, 836	535, 413, 007	412, 034, 067	417, 752, 697
H	548, 377, 594	520, 688, 854	446, 742, 473	378, 950, 050
I	572, 220, 316	493, 015, 494	415, 259, 267	382, 084, 372
des-rkGA	545, 289, 874	493, 015, 494	412, 034, 067	378, 950, 050

speed was faster than the other compared methods. As shown in Table 2, Z_1 was (402, 412, 188) when st-GA was used. It was (364, 869, 285) with rk-GA and (345, 695, 925) with des-rkGA. des-rkGA showed 16.41% and 5.55% improvements compared to st-GA and rk-GA, respectively. des-rkGA was confirmed to provide stable results because the standard deviation was only 3210 compared to 20, 338 with st-GA and 7780 with rk-GA. Table 3 shows the evolutive process of each island when popSize was 100 as in Table 2. The highlight shows the best solution for each generation, of which the best was with des-rkGA. In a prior experiment, the PC and PM for island E were the best combination. However, in the evolutive process for each island, many of the best solutions were produced at other islands. Table 4 shows the solution at each immigration rate, and the best solution by rk-GA is shown as reference. The experimental results show that if the immigration rate surpassed 50%, the results become bad. Figure 19 was created from Table 4; the best value was produced at 10% migration rate.

Table 4. Immigration Rate of des-rkGA

Immigration rate (%)	Z_1 (yen)
0	352, 543, 849
5	349, 209, 452
10	346, 035, 987
15	349, 468, 738
20	353, 121, 891
50	373, 634, 700
rk-GA	364, 769, 003

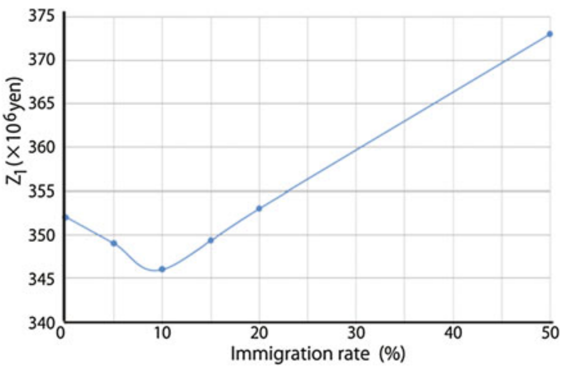


Fig. 19. Immigration Rate of des-rkGA

Table 5 summarizes the experimental results in terms of the des-rkGA improvement rate and standard deviation for Z_1 of maxGen when popSize

Table 5. Experimental results

	st-GA	rk-GA	des-rkGA
Z_1 (yen)	402, 412, 188	364, 869, 285	345, 695, 925
Improvement rate	16.41%	5.55%	-
standard deviation	20, 338	7, 780	3, 210

Table 6. Experimental results

	st-GA	rk-GA	des-rkGA	
			Each PC	Total PCs
maxGen	55.56	29.96	32.47	97.41
For 1d	35.78	19.98	20.45	61.35
Time to arrive at the maxGen value	55.56	11.07	5.74	17.22
Time to arrive at the rk-GA value	-	-	3.15	9.45

was 100. As shown in Table 2, there were differences when the number of generations was small.

In this experiment, we used three PCs of the same kind dual core AMD1212 2.0 GHz/2 MB; the memory size was 2 GB, and the development language was C#.

The computational time is shown in Table 6. We experimented with the test data for 90 days. When converted to 1 day, the computation time was 35.78, 19.98, and 20.45 s using st-GA, rk-GA, and des-rkGA, respectively. The average generation number when the solution arrived at the maxGen value was 685, 467, and 78 using st-GA, rk-GA, and des-rkGA, respectively. The generation time is shown in Table 6 as time to arrive at the maxGen value. The time was 55.56, 11.07, and 5.74 s using st-GA, rk-GA, and des-rkGA, respectively. There was no drastic improvement because the total CPU processing time was 9.45 s, but the CPU processing time could be distributed. Moreover, the average of the generation numbers when the des-rkGA value became better than maxGen of rk-GA was 165, and it took 3.15 s. We also confirmed the advantage of des-rkGA in terms of computer time. This was due to using parallel processing with 3 PCs of the same kind. We believe that the PC environment affects the solutions.

6 Conclusions

Using date of an actual automobile company in this study, we proposed random key-based genetic algorithm with distributed environment scheme (des-rkGA), for a multi-stage logistics system that calculates inventory values for many different cases. We proposed a logistics system that keeps the safety inventories only in DCs and that can cope with the location allocation problem. We performed numerical experiments with test data that were based on the disclosed data of a certain automobile company. The proposed des-rkGA showed improvements of 16.25% and 5.43% compared to the performances of st-GA and rk-GA, respectively. The des-rkGA also showed a small dispersion of solutions compared with the other methods. For future work, we intend to investigate the effectiveness of the logistics system in actual conditions while continuing to advance research and analysis of case studies.

Acknowledgements. This work is partly supported by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS): No. 15K00357.

References

1. Altiparmak F, Gen M, Lin L (2004) A priority-based genetic algorithm for supply chain design. Comput Ind Eng 13:22–25
2. Cauty R (2000) Genetic algorithms and engineering optimization. Wiley, Hoboken
3. Gen M, Cheng R (1997) Genetic algorithms and engineering design. Wiley, New York
4. Dobos I (2001) Production strategies under environmental constraints: continuous-time model with concave costs. Int J Prod Econ 71(1–3):323–330
5. Dobos I (2005) The effects of emission trading on production and inventories in the Arrow-Karlin model. Int J Prod Econ 93–94(1):301–308
6. Dobos I (2007) Tradable permits and production-inventory strategies of the firm. Int J Prod Econ 108(1–2):329–333
7. Gen M, Lin L (2009) Genetic algorithms. In: Wah B (ed) Wiley encyclopedia of computer science and engineering. Wiley, pp 1367–1381
8. Graves SC, Willems SP (2003) Supply chain design: safety stock placement and supply chain configuration. Handbooks Oper Res Manage Sci 11:95–132
9. Inoue H (2008) Utilization of genetic algorithm for transportation planning improvement in multi-objective logistics system. J Soc Plant Eng Jpn 19:252–259 (in Japanese)
10. Inoue H, Gen M (2012) A multistage logistics system design problem with inventory considering demand change by hybrid genetic algorithm. IEEJ Trans Electron Inf Syst 95(5):56–65
11. Jamrus T, Chien CF et al (2015) Multistage production distribution under uncertain demands with integrated discrete particle swarm optimization and extended priority-based hybrid genetic algorithm. Fuzzy Optim Decis Making 14(3):265–287
12. Lagodimos AG, Koukoumialos S (2008) Service performance of two-echelon supply chains under linear rationing. Int J Prod Econ 112(2):869–884
13. Lee JE, Chung KY et al (2015) A multi-objective hybrid genetic algorithm to minimize the total cost and delivery tardiness in a reverse logistics. Multimedia Tools Appl 74(20):9067–9085
14. Lin L, Gen M (2008) An effective evolutionary approach for bicriteria shortest path routing problems. IEEJ Trans Electron Inf Syst 128(3):416–423
15. Lin L, Gen M (2009) A random key-based genetic algorithm for agv dispatching in FMS. Int J Manufact Technol Manage 16(1):58–75
16. Mahadevan B, Pyke DF, Fleischmann M (2003) Periodic review, push inventory policies for remanufacturing. Soc Sci Electron Publ 151(3):536–551
17. Min H, Zhou G (2002) Supply chain modeling: past, present and future. Comput Ind Eng 43:231–249

18. Minner S (2003) Multiple-supplier inventory models in supply chain management: a review. Int J Prod Econ 81–82(02):265–279
19. Minner S, Silver EA (2005) Evaluation of two simple extreme transshipment strategies. Int J Prod Econ 93–94(1):1–11
20. Miranda PA, Garrido RA (2004) Incorporating inventory control decisions into a strategic distribution network design model with stochastic demand. Transp Res Part E Logistics Transp Rev 40(3):183–207
21. Miyata S, Kudo K et al (2003) Mass simulations based design approach and its environment: multi-objective optimization of diesel engine with distributed genetic algorithm using iSIGHT MOGADES and HIDECS. Parallel computational fluid dynamics, pp 499–506
22. Neungmatcha W, Sethanan K et al (2013) Adaptive genetic algorithm for solving sugarcane loading stations with multi-facility services problem. Comput Electron Agric 98(7):85–99
23. Porter ME (1985) Competitive advantage: creating and sustaining superior performance, pp 33–69
24. Rieksts BQ, Ventura JA, Herer YT (2009) Power-of-two policies for single-warehouse multi-retailer inventory systems with order frequency discounts. Comput Oper Res 36(7):2286–2294
25. Saul AT et al (1993) Numerical recipes in C. Cambridge University Press, Camridge, pp 293–295
26. Syarif A, Gen M (2003) Double spanning tree-based genetic algorithm for two stage transportation problem. Knowl Based Intell Eng Syst 7(4):214–221
27. Taniar D, Leung C et al (2008) High performance parallel database processing and grid databases. Wiley, New York
28. Thangam A, Uthayakumar R (2008) A two-level supply chain with partial backordering and approximated poisson demand. Eur J Oper Res 187(1):228–242
29. Zhou G, Min H, Gen M (2002) The balanced allocation of customers to multiple distribution centers in the supply chain network: a genetic algorithm approach. Comput Ind Eng 43:251–261

The Impact of Temporary Employment on Employees' Organizational Citizenship Behavior and Turnover Intention: The Moderating Effect of Organizational Identification

Xiaoye Qian[1], Qian Li[2(✉)], Qiong Wu[2], and Yilun Wu[3]

[1] Business School of Sichuan University, Chengdu 610064, People's Republic of China
[2] International Business School, Beijing Foreign Studies University, Beijing 100089, People's Republic of China
liq@bfsu.edu.cn
[3] School of Mechanical, Aerospace and Civil Engineering, The University of Manchester, Manchester M13 9PL, UK

Abstract. Temporary employment, compared with permanent employment, impacts on employees' attitudes and behaviors. This study provides empirical evidence on the relationships between temporary employment and employee's turnover intention and organizational citizenship behavior. Using a two-wave panel design, we collect data from a sample of 355 employees from 66 teams in four Chinese organizations. Our results show that, temporary employees have higher turnover intention than permanent employees, but in terms of employee's OCB, the difference between temporary and permanent employees is not significant. Further, we also explore the mechanism how employee'ss organizational identification as a moderator-for temporary employees high in organizational identification, the negative relationships between temporary employment and OCBs and the positive relationships between temporary employment and turnover intention would be both attenuated.

Keywords: Temporary employment · Organizational citizenship behavior · Turnover intention · Organizational identification

1 Introduction

To be competitive in a globalized economy, many organizations have been using nonstandard work arrangements to reconstruct their workforce to improve flexibility, to reduce employment-related costs, therefore improve competitiveness in product market [2]. One wide-adopted flexible employment practice conducted by Chinese firm is temporary employment. Kalleberg and his colleagues define standard work as "work done on a fixed schedule-usually full-time-at the

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_64

employer's place of business, under the employer's control, and with the mutual expectation of continued employment"; any arrangement that lacks one or more of these attributes is nonstandard [7]. There are many different forms of nonstandard work arrangements, such as temporary, contract, part-time and temporary agency work, to enhance organizational flexibility and reduce employment-related cost [7]. The temporary employment, as a most widely adopted form of nonpermanent employment, refers to employment featured with shorter length of contracts and lower expectations of continued employment.

According to the latest report by United Nations International Lab our Organization, in the United States, one in four employees worked part time in 2014, up from 19.6% in 2009. In 33 European countries, 12.3% employees were averagely on temporary contracts in 2014. In Asia, the portion of temporary workers are even higher, ranging from 24% in Philippines to 67% in Viet Nam. The percentage is sizeable in China, India, Indonesia and Malaysia.

In response to the growing use of nonstandard workers, more scholars are conducting empirical and theoretical research on this phenomenon. Much of the studies focus on the differences between standard and nonstandard employees in work-related attitudes, including satisfaction, commitment, loyalty, and in-role & extra-role behaviors, like organizational citizenship behavior, turnover and work performance [2,5,12,14].

Other than permanent employment, Contrast with the wide-acceptance in management practices, scholars have argued that there are quite a few drawbacks of temporary employment. Many studies reported that temporary jobs are connected with lower work status, characterized with low-wage, low-welfare and high-stress [4]. Scholars also observed that temporary employees often put in less effort than standard employees, hence alternative work arrangements increase the difficulty in human resource management [12].

Previous studies in U.S. and Euro examined the relationship between temporary employment and employee's work attitudes and performance. Christin and Linn [14] recruited a sample of 350 entry-level employees from 6 restaurants in United States, and found that part-time workers did show some organizational citizenship behavior (OCB) differences compared to their full time colleagues. Thorsteinson [5] did a meta-analysis to demonstrate that, full-time employees were more involved in their jobs than part-time employees and there was little difference between full-time and part-time employees in job satisfaction, organizational commitment, and turnover intention.

Compared with the abundant empirical studies in western countries, research on temporary employment has not attracted enough attention in China. To bridge the gap in the literature, we use a two-wave panel design study to examine whether the different work arrangements would influence employees' OCB and turn over intention. We have two major contributions to the literature. Firstly, this study is to investigate the relationship between temporary employment status and OCB and turnover intention in Chinese cultural context. Our research fills this gap in the literature and advances current understanding on blended workforce arrangement and employees' OCB and turnover intention.

Secondly, we test the hypothesized moderating effects of individuals' organizational identification on the relationships between work arrangements and employee's OCB and turnover intention. We propose that, the negative relationship between temporary employment and OCB will be attenuated, while the positive relationship between temporary employment status and turnover intention will be we propose that organizational identification play the role of moderating the relationship between temporary employment, OCB and turnover intention. The moderator is of great importance for us to understand how the effect works.

This paper is organized as follows. Section 2 presents concepts and hypotheses. Section 3 introduces the methodology. Section 4 summarizes the results and Sect. 5 concludes the paper.

2 Concepts and Hypotheses

2.1 Temporary Employment

During the twentieth century, the "normal" or "standard" work arrangement for most workers was that they were hired full-time and would work for the organization until they retired [14]. However, to deal with the changing internal and external environment, an increasing number of enterprises realize that they should not only boost productivity but also maintain flexibility. Therefore, many of them started to adopt a "non-standard" employment, hiring both "contingent" or "fixed-term contract" workers and standard employees [5].

The differences between "standard" and "contingent" employment arise from two fundamental criteria: the duration of their employment relationship and the level of HR investments [6]. While standard employees have sustained employment relationships and receive higher level of HR investments (e.g., economic incentives, training, empowerment opportunities), contingent ones have employment relationships of limited duration and generally receive lower level of HR investments [13].

2.2 Temporary Employment and Organizational Citizenship Behavior

There are some studies on the relationships between temporary employment and OCB. Van Dyne and Ang [3] tested their hypotheses based on a sample of 155 professional workers from 2 large service organizations, a bank and a hospital in Singapore. They found that the relationship between two kinds of commitment and psychological contracts and OCBs were stronger for temporary workers than standard employees. In contrast, Conway and Briner's analyses [1] showed that the relationships between psychological contract fulfillment and outcomes were rarely moderated by work status, suggesting that part-time employees would respond in a similar way as full-time employees to adjust their psychological contract. It seems that the relationship between temporary employment and OCB remains in dispute.

Thus, we propose the following hypothesis:

Hypothesis 1: Temporary employees are less likely to engage in organizational citizenship behavior, as compared to permanent employees.

2.3 Temporary Employment and Turnover Intention

Turnover intention can be measured by ideas of quitting, expectation of finding alternative employment, job search behavior and intention to quit. It is an important indicator to measure employees' emotional condition. If an employee has a high turnover intention, it is likely that he will not perform his job with all the effort.

Regarding existing studies on turnover intention, Peter, Jackofsky & Salter [11] found that there was differential predictability of turnover across full-time and part-time employees, suggesting that the two groups may have a different "psychology of work". Accordingly, we come up with the second hypotheses:

Hypothesis 2: Temporary employees are more likely to have turnover intention, as compared to permanent employees.

2.4 The Moderating Effect of Organizational Identification

Organizational identification is defined as a perceived oneness with an organization and the experience of the organization's successes and failures as one's own. The more people identify with an organization, the more the organization's values, norms, and interests are incorporated in the self-concept [8]. Knippenberg and Sleebos [8] found that organizational identification has a significant impact on turnover intention. Following this direction of this analysis, we propose the hypothesis below:

Hypothesis 3: Organizational identification will moderate the relationships between temporary employees and employee's OCB. Specifically, for employees high in organizational identification, the negative relationships between temporary employees and OCB will be attenuated.

Hypothesis 4: Organizational identification will moderate the relationships between temporary employees and employee's turnover intention. Specifically, for employees high in organizational identification, the positive relationships between temporary employees and turnover intention will be attenuated.

The framework of this research is shown in Fig. 1.

3 Methodology

3.1 Data

This study was conducted in 4 hospitals in Shaoxing City, Zhejiang Province of People's Republic of China (PRC), which are Lingzhi Town Hospital, Mashan

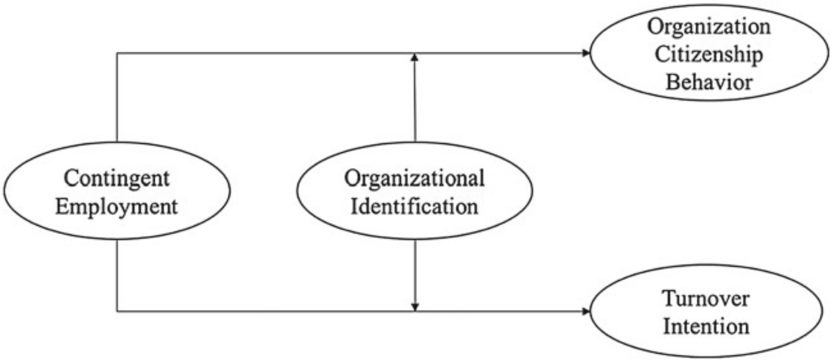


Fig. 1. Theoretical model of this study

Town Hospital, Sunduan Town Hospital and Fushan Community Hospital. All the hospitals are public community hospitals with more than 150 employees, providing service to people living in the town or the community.

The sample consists of both standard employees and temporary employees, including doctors, nurses and other employees like drivers and accountants who provide assistant service. Using information obtained from the leaders of 4 hospitals, we distributed questionnaires to each individuals separately. We finally got 209 valid questionnaires with a response rate of 80.2% (shown in Table 1).

Table 1. Summary of the sample

		N	%
Sample	Valid sample	209	57.89
	Invalid sample	152	42.11
	SUM	361	100

3.2 Measurement

Unless otherwise noted, we use a 5-point Likert scale ("strongly disagree" = 1 to "strongly agree" = 5) for all scales. Each scale's coefficient α is noted below in Table 5. Because the questionnaires are distributed in Chinese organizations, we follow the strict procedure of translating, back-translating and cultural adjustment of the original scales to generate a Chinese version.

(1) Employment Status
We collect the information of subordinates' employment status with a three-step procedure. Firstly, the supervisors are asked to report the subordinates' employment status two weeks before the survey. Then in the first wave survey, the subordinates report their own employment status. The survey question is "You are currently 1 'permanent employee', 2 'contract employee', or 3

Table 2. Organizational citizenship behavior used in this study

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
m56. I will attend functions that are not required but that help the hospitals image	1	2	3	4	5
m57. I will keep up with developments in the hospital	1	2	3	4	5
m58. I will defend the hospital when others criticize it	1	2	3	4	5
m59. I will show pride when representing the hospital in public	1	2	3	4	5
m60. I will offer ideas to improve the functioning of the hospital	1	2	3	4	5
m61. I will express loyalty toward the hospital	1	2	3	4	5
m62. I will take action to protect the hospital from potential problems	1	2	3	4	5
m63. I will demonstrate concern about the image of the hospital	1	2	3	4	5

'dispatched employee"'. Lastly, we identify the respondents' employment status through crosschecking the employment status information obtained from both parties and asking the HR department to verify inconsistent information. In the empirical analysis, the three different employment types are categorized into two: 1 is the standard employee, 0 is the temporary employee (contract employee and dispatched worker).

(2) Organizational Citizenship Behavior
Organizational citizenship behavior was measured by using an 8-item OCB scale developed by Organ [10]. Here we only used 8 items which are beneficial to the organization. Examples are: "I will attend functions that are not required but that help the hospital's image", "I will keep up with developments in the hospital" (Cronbach's α = .90). All the items of the OCB scale are listed in Table 2.

(3) Turnover Intention
Four items were used to assess each employee's turnover intention. Examples are: "I have frequent thoughts of leaving this hospital", "I often seek information

Table 3. Turnover intention items used in this study

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
m25. I have frequent thoughts of leaving this hospital	1	2	3	4	5
m26. I intend to find a new job	1	2	3	4	5
m27. I often seek information about other job possibilities	1	2	3	4	5
m28. I will quit my job soon	1	2	3	4	5

about other job opportunities" (Cronbach's $\alpha = .94$). All of items of the turnover intention scale are listed in Table 3.

(4) Organizational identification
Organizational identification was measured by using a 6-item scale developed by Mael and Ashforth [9]. Examples are: "When someone criticizes my team, it

Table 4. Organizational identification items used in this study

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
e24. When someone criticizes my team, it feels like a personal insult	1	2	3	4	5
e25. When someone praises my team, it feels like a personal compliment	1	2	3	4	5
e26. My teams successes are my success	1	2	3	4	5
e27. I am very interested in what others think about my team	1	2	3	4	5
e28. When I talk about my team, I usually say "we" rather than "they"	1	2	3	4	5
e29. If others criticized my team, I would feel embarrassed	1	2	3	4	5

feels like a personal insult", "I am very interested in what others think about my team" (Cronbach's $\alpha = .91$). All the items of the turnover intention scale are listed in Table 4.

(5) Control Variables
We controlled five demographic variables, which are age, gender, education background, organizational tenure, and employees' difference. Age and tenure were measured in years. Gender was measured as a dummy variable, 1 for male and 0 for female. As for education background, we used 1 to represent bachelors and above, 2 to represent technical school graduates, 3 to represent senior high school graduates and below. Treatment difference is employees' cognition on the organization's treatment differences between temporary workers and permanent workers on income level, welfare, uniform, training opportunities, etc.

4 Results

4.1 Descriptive Analysis

Tables 5 and 6 provide the descriptive statistics and correlations of the variables in this study. In the Table 5, the mean, standard deviation, minimum and maximum value of each variable are reported. As the figures in Table 5 shows, in a final sample ($N = 209$), 26.3% are male. The average age of the participants are 37.6 years old ($SD = 8.76$). The average organizational tenure is 11.7 years ($SD = 8.83$). As for education background, we used 1 to represent bachelors and above, 2 to represent technical school graduates, 3 to represent senior high school graduates and below. The average education background of participants is 1.41 ($SD = 0.69$).

Table 6 reports the correlation matrix of the research variables. As expected, independent variables of employment status are significantly correlated with organizational identification ($r = 0.17$), turnover intention ($r = -0.12$) and OCB ($r = 0.12$). The significance of correlation between employment status and the organizational identification, turnover intention and OCB indicate that permanent employees show higher organizational identification, less turnover intention and more OCB compared, as compared to temporary employees.

The Cronbach's alphas of major variables are listed in Table 6. The Cronbach's alphas of organizational identification, OCB and turnover intention's Cronbach's alphas are 0.91, 0.91 and 0.94. All the figures exceed the acceptable criteria, which is 0.70, which suggests that the validity of the scales chosen in this study is acceptable.

4.2 Confirmatory Factor Analysis

To further determine the factor structure of your data set, we conduct the confirmatory factor analysis of OCB, turnover intention and organizational identification. For OCB, the factor loadings are range from 0.60 to 0.85. For turnover

Table 5. Descriptive statistics of research variables

Variable	Number	Mean	Std. Dev.	Min	Max
Age	209	37.56938	8.755011	20	67
Edu	209	1.416268	0.689149	1	3
Gender	209	0.2631579	0.4414046	0	1
Tenure	209	11.70793	8.831362	0	46
Treatment difference	209	1.516373	0.358681	1	2.888889
ES	209	0.8373206	0.3699591	0	1
OI	209	3.844169	0.6415724	2	5
TUR	209	1.745135	0.9175642	1	5
OCB	209	3.551993	0.6366862	1	5

Note: Edu = education background; Gender is coded 0 = female, 1 = male; Tenure = working years; Treatment difference = the differences between permanent and temporary employees of how the organization treat them in uniform, wage, benefits, training and promotion, 5-likert scale; ES = Employment status, it is coded as 0 = contingent employee, 1 = regular employee; OI = organizational identification; TUR = turnover intention; OCB = organizational citizenship behavior.

Table 6. Correlations and cronbach's alpha of research variables

Variable	1	2	3	4	5	6	7	8	9
1. Age	7.1	-	-	-	-	-	-	-	-
2. Edu	0.47^{***}	1	-	-	-	-	-	-	-
3. Gender	0.27^{***}	0.128^{*}	1	-	-	-	-	-	-
4. Tenure	0.67^{***}	0.30^{***}	0.09	1	-	-	-	-	-
5. Treatment difference	0.03	0.14^{**}	0	-0.06	1	-	-	-	-
6. ES	-0.13^{**}	-0.53^{***}	-0.03	0.03	-0.32^{***}	1	-	-	-
7. OI	-0.06	-0.11	-0.15^{**}	-0.02	-0.08	0.17^{**}	-0.91	-	-
8. TUR	-0.20^{***}	-0.03	0.08	-0.23^{***}	-0.01	-0.12^{**}	-0.25^{***}	-0.94	-
9. OCB	0.02	-0.09	0.03	-0.05	-0.01	0.12^{*}	0.57^{***}	-0.35^{***}	-0.91

Note: $N = 209$; $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

intention, the factor loadings are from 0.82 to 0.95. For organizational identification, the factor loadings are from 0.67 to 0.90. The detailed factor loading of each item of OCB are listed in Table 7 as an illustration. Since component reliability of each items are higher than the criteria, i.e., 0.45. The measurement in this study is acceptable.

We then test the *Hypotheses 1* and *2* first. Then after we mean-center the variables, we use moderated regression analyses to test *Hypotheses 3* and *4*.

Table 7. CFA of organizational citizenship behavior

Standardized	Coef.	OIM Std. Err.	z	$P > \|z\|$	95% Conf.	Interval
m56←OCB	0.598	0.049	12.28	0	0.502	0.693
_cons	4.19	0.216	19.37	0	3.766	4.614
m57←OCB	0.732	0.036	20.27	0	0.661	0.803
_cons	4.265	0.22	19.41	0	3.834	4.695
m58←OCB	0.695	0.04	17.52	0	0.617	0.773
_cons	4.222	0.218	19.39	0	3.796	4.649
m59←OCB	0.758	0.034	22.54	0	0.692	0.824
_cons	4.471	0.229	19.49	0	4.022	4.921
m60←OCB	0.685	0.04	16.98	0	0.606	0.764
_cons	3.715	0.194	19.11	0	3.334	4.096
m61←OCB	0.846	0.024	34.88	0	0.799	0.894
_cons	4.375	0.225	19.45	0	3.934	4.816
m62←OCB	0.762	0.033	22.81	0	0.696	0.827
_cons	4.834	0.246	19.62	0	4.351	5.317
m63←OCB	0.812	0.028	28.53	0	0.756	0.868
_cons	4.697	0.24	19.58	0	4.227	5.165

Note: LR test of model vs. saturated: chi2(20) = 107.71, Prob > chi2 = 0.0000

4.3 Employment Status and OCB, Turnover Intention

The results of OLS regression are presented in Table 8. In Model 1, we put in 5 control variables, including age, education background, gender, tenure and participants' difference. None of these five control variables has significant impact on employees' OCB. When the independent variable, "temporary employment", is added into the regression in Model 2, the results show that permanent employment has a positive impact on OCB but the impact is insignificant ($\beta = 0.202$, $p > 0.1$). In Model 3, we remove the control variable of education background, then we find that permanent employment has a significant positive impact of OCB ($\beta = 0.251$, $p < 0.1$). These results suggest that if education background is not considered, standard employees show more organizational citizenship behavior than temporary employees.

In Model 4, the dependent variable is turnover intention. We also put in 5 control variables, including age, education background, gender, tenure and participants'difference. Results show that age ($\beta = -0.018$, $p < 0.1$) has a negative impact on turnover intention while gender ($\beta = 0.272$, $p < 0.1$) has a positive impact on turnover intention. When the independent variable, "temporary employment", is added into the Model 5, results show that permanent employment has a significant negative impact on turnover intention ($\beta = -0.402$, $p < 0.1$). Therefore, *Hypothesis 2* is supported.

Table 8. Regressions results

Variable	Dependent variable				
	OCB			Turnover intention	
	M1	M2	M3	M4	M5
Age	0.108	0.010	0.009	−0.018	−0.018
Edu	−0.115	−0.055		0.108	−0.116
Gender	0.028	0.025	0.024	0.273	0.279
Tenure	−0.008	−0.010	−0.010	−0.156	−0.012
Difference	−0.100	0.038	0.040	−0.638	−0.160
ES		0.201	0.251**		−0.402**
R^2	0.020	0.029	0.027	0.077	0.090
ΔR^2	−0.004	0.000	0.003	0.054	0.070
F	0.830	1.000	0.636	3.380	3.470

Note: $N = 209$; $^{**}p < 0.1$, $^{**}p < 0.05$, $^{**}p < 0.01$.

4.4 The Moderating Effect of Organizational Identification

We also conduct OLS regression to examine *Hypothesis 3*. The results are reported in Model 6 of Table 9. In this model, besides the control variables, we introduce the independent variable Employment status (ES), the moderator organizational identification (OI) and the interactions of $ES \times OI$ (both variables are mean-centralized). Results show that the interactions of $ES \times OI$ has a positive impact on employees' OCB, but it is not significant ($\beta = 0.178$, $p > 0.1$). Therefore, the *Hypothesis 3* is not supported.

We follow the same steps to test *Hypothesis 4*. The results are demonstrated in Model 7 of Table 9. In this model, besides the control variables, we introduce the independent variable Employment status (ES), the moderator organizational identification (OI) and the interactions of $ES \times OI$ (both variables are mean-centralized). We find that the interactions of $ES \times OI$ has a significant positive impact on employees' turnover intention ($\beta = 0.638$, $p < 0.1$). Therefore, our *Hypothesis 4* is supported.

5 Conclusion

In this study, we investigate the relationships between temporary employment, employees' OCB and turnover intention. We also propose that employee's organizational identification is a moderator in the above relationships. Results show that if education is not controlled, standard employees show more OCB than temporary employees. Compared with temporary employees, standard employees show less turnover intention. Organizational identification positively moderates the relationship between temporary employees and turnover intention, indicating that if a standard employee has a higher organizational identification, the impact of employment form on turnover intention would be much stronger.

Table 9. Moderating Effect

Variable	Dependent variable	
	OCB M6	Turnover intention M7
Age	0.009	-0.019^{**}
Edu	-0.048	0.001
Gender	0.144^{**}	0.202
Tenure	-0.008	-0.012
Difference	-0.074	-0.149
OI	0.561	-0.385
$EF \times OI$	0.178	0.638^{**}
R^2	0.35	0.159
ΔR^2	0.324	0.126
F	13.45	4.74

Note: $N = 209$; $^{**}p < 0.1$, $^{**}p < 0.05$, $^{**}p < 0.01$.

Our research makes two contributions to current literature. First, we have examined a conceptual model in cross cultural context. Although several studies in west have investigated the relationships between temporary employment and in-role/out-role performance, we do not know whether their results are consistent in Chinese context. Therefore, it is theoretically important for us to investigate these relationships in China. Second, our research demonstrates the important role of organizational identification. More specifically, we contribute to the literatures by examining and confirming organizational identification as a moderator in the above relationships.

Our paper also provides some managerial implications. To maximize temporary employees' work efforts, it's vital to increase their job satisfaction by improving their salary and welfare. Meanwhile, human resource department need to take effective actions to promote enterprise culture among temporary employees, thus raising their organizational identification and loyalty to work.

This paper also has some limitations. First, all data were collected within a single industry, which decreases external validity. Second, we use the self-reported data for the analysis, which may raise the possible common method bias. We recommend that future research could collect the data from different resources. Whether the findings of our study apply to workers in other settings or to other types of workers, should be examined by future studies.

Acknowledgements. The authors thank for the support by the National Science Foundation of China (Grant No. 71402108), the Humanities and Social Sciences Foundation of the Ministry of Education (Grant No. 14YJC630103), the Fundamental Research Funds for the Central Universities (No. 2016JJ019) and Young Faculty Research Fund of Beijing Foreign Studies University (2016JT003).

References

1. Conway N, Briner RB (2002) Full-time versus part-time employees: understanding the links between work status, the psychological contract, and attitudes. J Vocat Behav 61(2):279–301
2. Davisblake A, Broschak JP, George E (2003) Happy together? How using non-standard workers affects exit, voice, and loyalty among standard employees. Acad Manage J 46(4):475–485
3. Dyne LV, Ang S (1998) Organizational citizenship behavior of contingent workers in singapore. Acad Manage J 41(6):692–703
4. Johnson CD, Messe LA, Crano WD (1984) Predicting job performance of low income workers: the work opinion questionnaire. Pers Psychol 37(2):291–299
5. Kalleberg AL (2000) Nonstandard employment relations: part-time, temporary and contract work. Ann Rev Sociol 26:341–365
6. Kalleberg AL (2001) Organizing flexibility: the flexible firm in a new century. Br J Ind Relat 39(4):479–504
7. Kalleberg AL, Reskin BF, Hudson K (2000) Bad jobs in America: standard and nonstandard employment relations and job quality in the United States. Am Sociol Rev 65(2):256–278
8. Knippenberg DV, Sleebos E (2006) Organizational identification versus organizational commitment: self-definition, social exchange, and job attitudes. J Organ Behav 27(5):571–584
9. Mael F, Ashforth BE (1992) Alumni and their alma mater: a partial test of the reformulated model of organizational identification. J Organ Behav 13(2):103–123
10. Organ DW (1988) Organizational citizenship behavior: the good soldier syndrome. Adm Sci Q 41(6):692–703
11. Peters LH, Jackofsky EF, Salter JR (1981) Predicting turnover: a comparison of part-time and full-time employees. J Organ Behav 2(2):89–98
12. Stamper CL, Dyne LV (2003) Organizational citizenship: a comparison between part-time and full-time service employees. Cornell Hotel Restaurant Adm Q 44(1):33–42
13. Stirpe L, Trullen J, Bonache J (2013) Factors helping the hr function gain greater acceptance for its proposals and innovations: evidence from spain. Int J Hum Res Manage 24(20):3794–3811
14. Thorsteinson TJ (2003) Job attitudes of part-time vs. full-time workers: a meta-analytic review. J Occup Organ Psychol 76(2):151–177

Integration of Sound and Image Data for Detection of Sleep Apnea

Takehiro Kasahara[1,2](✉), Kota Nomura[3], Yoshihiro Ueda[2], Yuji Yonezawa[2], Masatoshi Saito[4], Hirohisa Toga[4], Yuki Fujimoto[4], Koji Kojima[4], Haruhiko Kimura[1], and Hidetaka Nambo[1]

[1] Kanazawa University, Kakuma, Kanazawa, Ishikawa 920-1192, Japan
kasaharat@irii.jp
[2] Industrial Research Institute of Ishikawa, 2-1 Kuratsuki, Kanazawa, Ishikawa 920-8203, Japan
[3] Akatsuki Inc., 8F Oak Meguro, 2-13-30 Kamiosaki, Shinagawa-ku, Tokyo 141-0021, Japan
[4] Kanazawa Medical University, 1-1 Daigaku, Uchinada, Kahoku, Ishikawa 920-0293, Japan

Abstract. To detect sleep apnea syndrome (SAS), we propose a method that integrates sound and image processing results using overnight sound and video data. In sound processing, snoring sounds are initially extracted using an unsupervised method. The zero-cross ratio (ZCR) and power (PWR) are calculated for the snoring sounds, and the total validations (TV) for ZCR and PWR are calculated using the sound data. Then, a support vector machine (SVM) is used to classify SAS and non-SAS subjects using TV values. In image processing, inter-frame absolute difference values are calculated every 1 s from video data, and multiscale entropy (MSE) is calculated using the time series data. Then, an SVM is employed to classify SAS and non-SAS subjects using two MSE values selected based on the Bhattacharyya distance. To integrate sound and image data, we focused on the number of detected snores. When two results do not match and the number of extracted snores is greater than five per hour, the sound processing result is adopted. When the number of snores is less than five per hour, the image processing result is adopted. The proposed integrated method demonstrates better accuracy than that obtained using only sound or image processing results.

Keywords: Sleep apnea · Hypopnea · Integration of image and sound processing · Healthcare services

1 Introduction

Sleep apnea syndrome (SAS) is a disorder in which breathing stops during sleep. SAS can cause sleepiness during the day, which may lead to traffic accidents. In addition, SAS induces a serious circulatory organ disease [6]. SAS symptoms

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_65

occur in approximately 4% of men and 2% of women. If people without subjective symptoms are included, this increases to 24% of men and 9% of woman [11], which is a significant portion of the population. In Japan, the prevalence of SAS is thought to be the same or more than in the United States [9]. Apnea is a temporary cessation of breathing for more than 10 s while sleeping, and hypopnea is an abnormally low ventilation, i.e., less than 50% of normal, lasting more than 10s. The apnea hypopnea index (AHI) represents the occurrence frequency per hour. Apnea is diagnosed when the AHI is five or more. Apnea and hypopnea can be classified as obstructive, central, or mixed. Simple noninvasive SAS diagnostic tools include pulse oximeters and apnomonitors. More reliable diagnoses can be obtained by polysomnography (PSG). PSG is performed in a hospital and requires multiple sensors. To collect data, the subject is required to stay in hospital overnight. Various techniques to detect SAS have been proposed. These techniques use sound data [2], video data [4], and a combination of depth video and sound [10]. Previously, we proposed a method that uses microphones and cameras, which are widely used noncontact sensors [7].

2 Methods

In this study, we have improved the previously proposed method. We used PSG data and evaluated the accuracy of the improved method to classify SAS and non-SAS more precisely by integrating the results of sound and image processing.

2.1 Data Acquisition

Overnight PSG data were collected from the sleep studies of 55 patients (nine females, 61.2 ± 12.4 years) who were suspected to suffer from SAS and referred to the Kanazawa Medical University (Ishikawa, Japan). The subjects provided written consent to participate in the sleep studies, which were conducted in 2014. Four subjects were diagnosed as normal and 51 were diagnosed with different levels of SAS. To obtain more non-SAS subject data, overnight pulse oximeter data were collected from 24 employees of the Industrial Research Institute of Ishikawa or CosmoSummit Co., Ltd. in 2014–2015. These employees, none of whom complained about sleep apnea symptoms, also provided written consent, took data according to a data acquisition manual themselves. From the analysis of the pulse oximeter data, three subject datasets were considered to indicate SAS, and 24 subject datasets were considered to indicate non-SAS. Subject information, such as age, gender, height, and weight, was also collected (Table 1).

The overnight sounds of the participants were recorded using a microphone (SONY ECM-360) placed near the subject's head. The sampling rate was 11025 Hz at 16-bit resolution. Overnight video of the participants was recorded using a camera (IDS UI-1220LE-M-GL) with a lens (TAMRON12VM412ASIR) that is sensitive to infrared light. The camera was placed at the side of the bed. An example video frame is shown in Fig. 1:

Table 1. Overview of subject data

Parameter	SAS subjects: 44 male, 10 female		Non-SAS subjects: 26 male, 2 female	
	Mean $\pm$ σ	Range	Mean $\pm$ σ	Range
Age	59.4 ± 14.2	26–88	38.5 ± 9.9	23–66
BMI	25.8 ± 3.6	17.2–33.4	22.8 ± 3.2	17.8–30.2
AHI	38.8 ± 18.8	5.5–76.0	2.6 ± 1.1	0.6–4.7

Fig. 1. Example video frame

2.2 Sound Processing

Since various sounds, e.g., subject and bedding movements, and coughs, are included in the recorded sound data, we used an unsupervised method proposed by Azarbarzin et al. [1] to extract snore sounds. The unsupervised snoring sound extraction was performed as follows. After applying a band pass filter (BPF, 150–5000 Hz) to the recorded data, a section in which the volume is greater than a predefined threshold continues for 0.4 to 2.0 s was extracted as an episode. Extracted episodes include snoring episodes and episodes due to noise other than snoring. For one episode, 50% overlap is multiplied by a 50-ms window, and short-time Fourier transform (STFT) was performed. The STFT results for an episode were averaged, and the spectral intensities in the range 0–5000 Hz (in practice, BPF in the range 150 to 5000 Hz) were summed every 500 Hz to create a 10-dimensional vector. Principal component analysis was performed using the 10-dimensional vector of all episodes. We reduced the dimensionality of the feature space to a two-dimensional vector, and classification was performed using the Fuzzy c-Means clustering method. In this study, we adopted the number three classes employed in Azarbarzin's method. Among the three classes, clusters with a greater number of episodes were considered snoring clusters, and episodes in this cluster were considered snoring episodes.

For snoring episodes extracted by unsupervised methods, SAS and non-SAS were classified using the method proposed by Azarbarzin et al. [2]. The zero-crossing rate (ZCR) and power value (PWR) were calculated for the extracted snoring episodes. Snoring sounds are highly variable relative to the severity of SAS. As severity increases, characteristics, such as frequency, will vary; thus, ZCR and PWR were used to calculate the amount of change (total validation norm, TV), i.e., TVzcr and TVpwr. The calculation of TVi is shown in Eq. (1). These two values are output as a two-dimensional feature amount for the recorded data from a single night. However, TVzcr and TVpwr output 0 when no snoring sounds were detected. The two-dimensional features are used for support vector machine (SVM) classification. A radial basis function kernel (RBF) was used with the SVM, and parameters C and were determined by grid search.

$$TV_i = \frac{1}{N}\sum_{k=1}^{N-1} |f_i(k+1) - f_i(k)|. \tag{1}$$

2.3 Image Processing

The video recordings had a duration of approximately 6 hours and were recorded in MJPEG format at a resolution of 752×480 and an average frame rate of 5 frames per second (fps). Considering that the fastest physiologically important body movements associated with SAS during sleep are respiratory movements and the respiratory rate for adults at rest is on average 12–18 breaths per minute, the frame rate was reduced to 1 fps. The overall motion signal was derived by calculating the mean absolute pixel difference between successive frames.

Multiscale entropy (MSE) analysis was performed using the obtained signal. MSE analysis is an analytical method proposed by Costa et al. [3], which is highly useful for physiological indices, especially heart rate variability. Generally, for a signal with high periodicity, the values obtained are lower, and the values obtained for waveforms with lower periodicity are higher. Therefore, by adapting the MSE analysis to the respiratory motion signal, the disorder degree of the periodicity of the respiratory signal was evaluated and used as a classification parameter.

MSE analysis is primarily performed in two steps. First, a coarse grain process averages the original data $x = \{x_1, x_2, \cdots, x_n\}$ by t to obtain data $y^{(t)} = \{y_1, y_2, \cdots, y_{N(=n/t)}\}$. Next, entropy is measured for each coarse-grained time series corresponding to t. Different measures of entropy can be employed for this step. Here, sample entropy proposed by Richman and Moorman [8] was used as the entropy metric. This metric takes two parameters, i.e., pattern length m and similarity criterion r, which is the tolerance for accepting pattern matches. r is a positive real value and is typically chosen between 10% to 25% of the standard deviation of the time series. Two patterns of length m match if each point in the first pattern is within distance r from the corresponding point in the second pattern. The distance between two vectors is defined as the maximum absolute difference between the components of those two vectors. A vector $x_m(i) = \{x_i, \cdots, x_{(i+m)-1}\}$ where $1 \leq i \leq N - m$ is defined as a pattern of length m, Note that i can only go up to $N - m$ to ensure that vector x_{m+1} of length $m + 1$ is also defined. $n_i^m(r)$ is defined as the number of vectors $x_m(j)$ that have distance less than r with $x_m(i), (i \neq j)$ to exclude self matches. $U_i^m(r) = n_i^m(r)/(N - m)$ is the probability that the distance between vector $x_m(i)$ and any other vector $x_m(j)$ is less than r. $U^m(r)$ is the probability of any two vectors of length m being within distance r of each other.

Therefore, the probability $U^m(r)$ is expressed as

$$U^m(r) = 1/(N - m) \sum_{i=1}^{N-m} U_i^m(r). \tag{2}$$

The sample entropy is defined as

$$H_{SE}(m, r) = \lim_{N \to \infty} - \ln \frac{U^{m+1}(r)}{U^m(r)}. \tag{3}$$

This is estimated by the following statistic

$$H_{SE}(m, r, N) = - \ln \frac{U^{m+1}(r)}{U^m(r)}. \tag{4}$$

The sample entropy was calculated for each subject's motion signal over scale factors $t \in T = \{1, \cdots, 60\}$, pattern lengths $m \in M = \{2, 3, 4, 5, 6\}$, and similarity criteria $r \in R = \{0.10, 0.15, 0.20, 0.25\}$. Therefore, a total of 1200 sets of entropy values were calculated for all subjects.

The Bhattacharyya distance [5] is a measure of similarity of two continuous or discrete probability distributions. In classification, it is used as a measure of separability between two classes. The distance is calculated from the Bhattacharyya coefficient ρ, which is a measure of the amount of overlap between two probability distributions. For discrete distributions p and q over a domain X, the Bhattacharyya coefficient is defined as follows

$$\rho = \sum_{x \in X} \sqrt{p(x)q(x)}, \tag{5}$$

where $0 \leq \rho \leq 1$. The Bhattacharyya distance Δ_B is defined as follows

$$\Delta_B = - \ln(\rho), \tag{6}$$

such that $0 \leq \Delta_B \leq \infty$. Low ρ values or high Δ_B values indicate little overlap between p and q. The 1200 combinations of MSE parameters were analyzed and ranked based on the distance between their corresponding sample entropy distributions of SAS and non-SAS subjects. Larger Bhattacharyya distance between the two distributions indicates that the sample entropy values are better feature candidates for classification.

An SVM with an RBF kernel was used for classification of SAS and non-SAS subjects. The features used in the classification were the sample entropy values, and parameters C and σ were determined by grid search.

2.4 Integration of Sound and Image Processing Results

We performed final classification from the results obtained by image processing and sound processing. If both results match, the result can be the final classification result. On the other hand, when each result differs, either result is adopted by considering the reliability of the acquired data. In this study, we considered the snoring number extracted from the sound data. When the number of extracted snores was greater than five per hour, the sound processing result was adopted. When the number of extracted snores was less than five per hour, the image processing result was adopted. In the data acquisition environment, the image data could not be acquired successfully because the degree of installation difficulty of the camera is greater than that of the microphone, and the distance is too close or the angle does not match. Therefore, when the extracted snore number is detected as 5 per hour or more, the sound processing result was adopted. In addition, several cases were confirmed where snoring did not occur even before and after the occurrence of apnea, and such cases could not be distinguished from cases where recording failed; thus, if snoring was not detected, the image processing result was adopted.

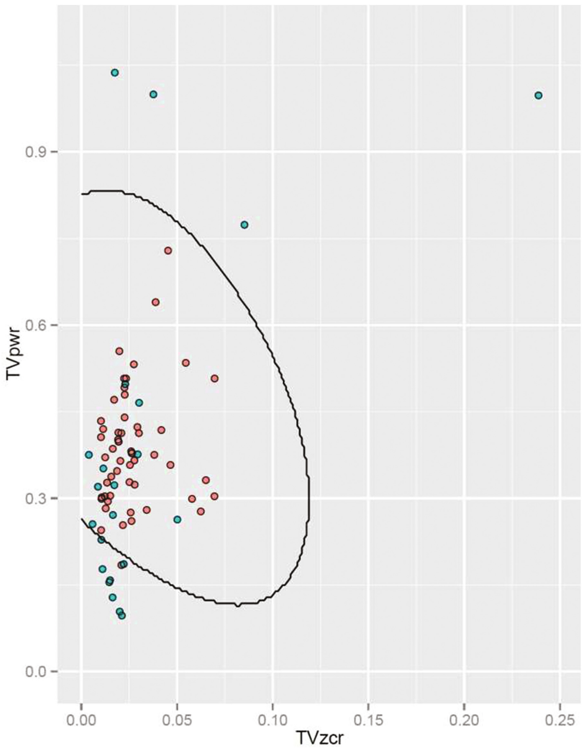


Fig. 2. SVM result in sound processing using all data

3 Results

In sound processing, non-SAS data with a small snore count were greatly deviated from the origin when plotted on $2D$ feature coordinate graph, and learning in the SVM did not provide proper results. The SVM result obtained using all data is shown in Fig. 2. Since snoring was not detected sufficiently, data with than five snores per hour were assumed to be non-SAS, and data classified by the SVM for the remaining data were taken as the sound processing result. The SVM result using data with less than five snores per hour is shown in Fig. 3:

In image processing, MSE analysis was performed to calculate the Bhattacharya distance, and the parameters sorted in descending order are shown in Table 2. The top of 2 parameter sets of $(m, r, t) = (4, 0.10, 30), (2, 0.25, 34)$ were adopted for SVM. The sample entropy when $(m, r) = (4, 0.10)$ is shown in Fig. 4. The result of SVM is shown in Fig. 5.

Accuracy was calculated by leave-one-out cross validation. With only image processing, the accuracy was 81.7%, sensitivity was 81.5%, and specificity was 82.1%. With only sound processing, the accuracy was 85.4%, sensitivity was 85.2%, and specificity was 85.7%. These results show that sound processing

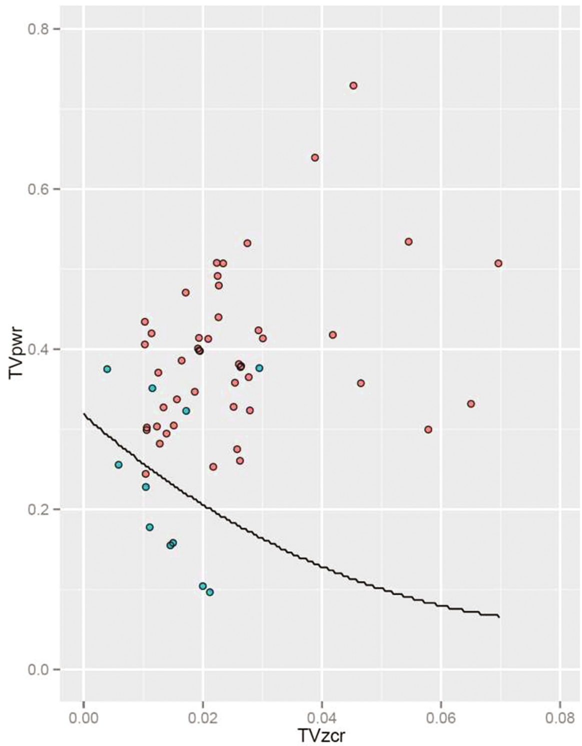


Fig. 3. SVM result in sound processing with more than five snores per hour

Table 2. Sample entropy parameters of the highest bhattacharyya dist

Order	Bhattacharyya dist	m	r	t
1	0.7140	4	0.10	30
2	0.7135	2	0.25	34
3	0.7119	2	0.25	44
4	0.7100	3	0.10	38
5	0.7074	2	0.10	49

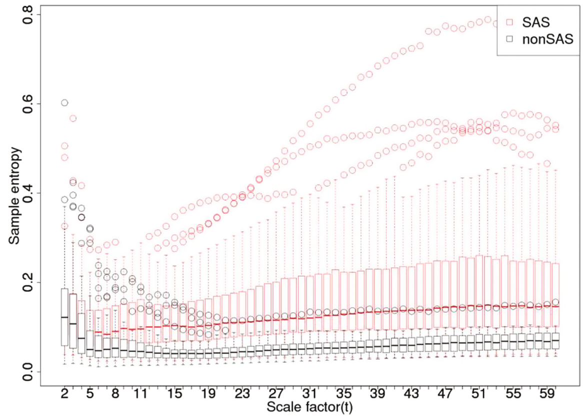


Fig. 4. Sample entropy when $(m, r) = (4, 0.10)$

outperformed image processing. In the integrated result, the accuracy was 89.0%, sensitivity was 92.6%, and specificity was 75.0%, which shows a decrease in specificity compared to the sound processing results; however, improvements to accuracy and sensitivity were observed. Table 3 compares classification results and the final results for image and sound processing.

Overnight PSG was used to obtain the number of apnea events and the number of hypopnea events. Subjects were classified as apnea or hypopnea type by comparing these events. In the sound processing results, there were eight false negatives out of 54 positive cases; however, in the integrated results, false negatives were reduced to four. In this case, among the four subjects that changed from false negatives to true positives, three subjects were hypopnea type. Of the 19 apnea type subjects, only one was mistaken as a false negative with sound processing. Thus, sound processing is effective for detecting apnea type SAS but less effective for hypopnea type SAS. However, by integrating the sound and image processing results, it is possible to reduce errors by half.

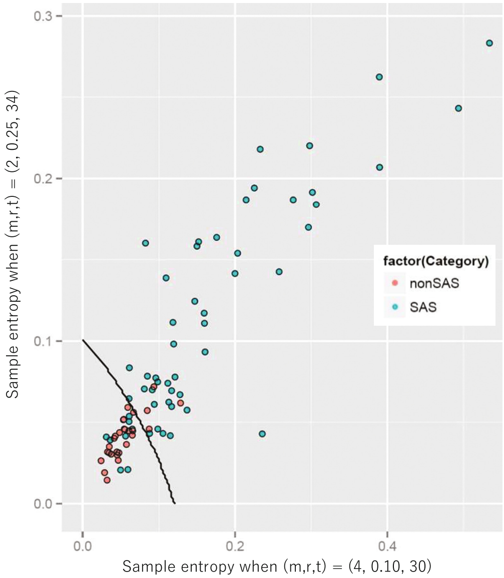


Fig. 5. SVM result in image processing using all data

Table 3. Individual results

Process	Image	Sound	Integrated
Accuracy (%)	81.7	85.4	89.0
Sensitivity (%)	81.5	85.2	92.6
Specificity (%)	82.1	85.7	75.0
Positive predictive value (%)	89.8	92.0	87.7
Negative predictive value (%)	69.7	75.0	84.0

4 Discussion and Conclusion

Screening methods that can easily detect SAS are expected. Therefore, we analyzed sound data and image data obtained using microphones and cameras, which are inexpensive and familiar sensors. We have proposed a screening method to classify SAS and non-SAS subjects. In this study, we focused on the number of snores per hour. When the number of extracted snores was large, we used sound processing due to good accuracy. However, low accuracy was obtained with sound processing when the number of snores was small; thus, we proposed a method using image processing. The experimental results demonstrate that the proposed integrated method (sound and image data) can screen at high accuracy compared to screening with only sound or image data. Furthermore, the proposed method is effective for screening hypopnea SAS.

Microphones and cameras are standard sensors on smartphones; therefore, it can be expected that SAS screening can be performed easily using smartphones by applying the results of this study.

References

1. Azarbarzin A, Moussavi Z (2011) Automatic and unsupervised snore sound extraction from respiratory sound signals. IEEE Trans Biomed Eng 58:1156–1162
2. Azarbarzin A, Moussavi Z (2012) Snoring sounds variability as a signature of obstructive sleep apnea. Med Eng Phys 27:479–485
3. Costa M, Goldberger A, Peng C (2002) Multiscale entropy analysis of complex physiologic time series. Phys Rev Lett 89:068102
4. Gederi E, Clifford G (2012) Fusion of image and signal processing for the detection of obstructive sleep apnea. In: 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics, pp 890–893
5. Kailath T (1967) The divergence and bhattacharyya distance measures in signal selection, communication technology. IEEE Trans Commun Technol 15:52–60
6. Kario K (2009) Obstructive sleep apnea syndrome and hypertension: ambulatory blood pressure. Hypertens Res 32:428–432
7. Nomuara K, Kasahara T et al (2015) Development of a system for detecting obstructive sleep apnea by combining image and sound processing. J Nanjing Univ Aeronaut Astronaut 27:12–20
8. Moorman Richman J JS (2000) Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circulatory Physiol 278:H2039–H2049
9. Tanigawa T, Tachibana N et al (2004) Relationship between sleep-disordered breathing and blood pressure levels in community-based samples of japanese men. Hypertens Res 17:479–484
10. Yang C, Cheung G et al (2016) Sleep apnea detection via depth video & audio feature learning. IEEE Trans Multimedia 1–5
11. Young T, Palta M et al (1993) The occurrence of sleep-disordered breathing among middle-aged adults. New Engl J Med 17:1230–1235

Pricing Strategy Study on Product Crowdfunding

Min Luo[1,2,3], Yong Zhang[4](✉), Huanmei Zhang[4], and Xia Li[4]

[1] Guanghua School of Management, Peking University, Beijing 100871, People's Republic of China
[2] Mathematics Department, Chengdu University of Information Technology, Chengdu 610225, People's Republic of China
[3] Uncertainty Decision-Making Laboratory, Sichuan University, Chengdu 610064, People's Republic of China
[4] School of Mathematical Sciences, Beijing Normal University, Beijing 100875, People's Republic of China
zhangyong@bnu.edu.cn

Abstract. The pricing strategy of crowdfunding has always been its highlights and difficulties. A pricing strategy formula concerning with the price discrimination is obtained through the in-depth studies of the price discrimination theory and the mechanism of crowdfunding pricing, considering the crowdfunding financial limits which constitute constraint conditions. This article carries out a study on the effect of advertising profit of crowdfunding and founds the profit function. Furthermore, it studies the model of crowdfunding based on the 1P theory to solve this issue. Finally, the article provides a corresponding pricing strategy on product crowdfunding based on the above studies. It is expected to contribute to the field of theoretical research on crowdfunding and the pricing strategy through this article, and provide inspiration and reference for enterprises and entrepreneurs when they make the pricing strategy decisions.

Keywords: Crowdfunding · Pricing strategy · Price discriminate · Profit

1 Introduction

Product crowdfunding refers that the investors put their money into developing a product or service in accordance with the fundraisers; when the product or service begins to be pre-sold or has the conditions for external sales, the fundraisers will provide the developed product or service for the investors. Product crowdfunding needs a third party to be supported by the public. The relevant information of the products, services and producers can be obtained through the platforms, which would save transaction cost while realizing a certain advertising effect.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_66

The consumers participating in the crowdfunding have strong interests in the project products when these are in the creating or initial stage, willing to provide funds to make the products available as soon as possible, and the price paid for each unit may not be lower than that of the product later on. It is an external direct financing that a project can be conducted smoothly when the crowdfunding reaches the planned fixed amount, which can reduce the difficulty of financing for the companies and can predict the market's reflection and selling through probing the participating action.

It is necessary to establish an adequate pricing and feedback module for a successful project. In general, in order to attract the public to support the participation of supporters, the project launchers will announce the prices of the products and a combination of preferential prices on the platform pre-sale in advance to the market. In addition, the use of the price discrimination strategy can tell the people who have different needs apart to sell. Grasping the psychological expectations of consumers and taking differentiated pricing not only can promote the products and return the funds ahead of time, but also can maximize the corporate profits.

2 Literature Review

For the platforms of the product crowd funding, pre-sale is generally the main way, which functions to find prices well and can help the project launchers or entrepreneurs to create some favorable conditions for price discrimination. Pigou [8] defines that the price discrimination refers to different prices charged by firms for different markets, consumers or different numbers of purchases, and he also puts forward a three-level price discrimination theory. As a kind of theory, price discrimination belongs to the category of the pricing strategy without judging. In the industries with more competition, price discrimination is widely used in a variety of flexible forms. It is an effective pricing strategy, not only helping to enhance the competitiveness of enterprises to achieve its business objectives, but also adapting to the psychological differences of consumers to meet the needs at multiple levels. Rochet and Thanassoulis [10] declares that the price discrimination is an economic phenomenon and a trading strategy, which is popularly used by the sellers; and according to economic thinking, the implementation of the price discrimination strategy for the sellers is positive in terms of the welfare and can thus expand their own profits.

Nocke et al. [7] studied the relations between the product crowd funding and the price discrimination in the context of asymmetric information. In an intertemporal setting in which the individual uncertainty is resolved over time, advance-purchase discounts can serve to the price discrimination between consumers with different expected valuations for the product. Consumers with a high expected valuation will purchase the product before learning the actual valuation at the offered advance-purchase discount; consumers with a low expected valuation will wait and purchase the goods at the regular price only in the occasion where their realized valuation is high. The consumers with higher expectations of participation in the product quality in the case of unknown product quality should enjoy discounts and have the lower prices than those for the regular customers. Enterprises have to target at the customers with different expectations. Belleflamme et al. [11] isolated some important features of crowdfunding on the basis of a unique, hand-collected dataset and proposed a model of crowdfunding that encompasses several of these key features. By constructing the profit function, they obtained the pricing expression of the product and the ownership of the product. On this basis, the enterprise financing limit is given, and the uncertainty of the product quality and information are elaborated. Hu et al. [2] and Ping [9] maximized the profits of producers as the objective function optimization model to analyze the pricing mechanism of the crowdfunding products, drawing the conclusion that the model will help to open up the initial public offering of the product market, and the profits of manufacturers can be expanded with the same total sales.

Lawton and Marom [5] pointed out that crowdfunding is not only a new source of funding for projects, but also a way to quickly attract the attention of users, form early user communities; and because of the large number of media reports and discussions, pre-release has become a common corporate marketing strategy, so crowdfunding is still a very effective marketing tool. Based on the choices of consumers, Bayus [1] applied the model of maximizing their own profits to jointly determine the time choice and pricing decision from the consumer utility function. Zhang et al. [14] and Luo et al. [6] constructed a two-stage pricing model based on the impact of IWOM on customers' perceived value to the product. IWOM can change the cognitive value of customers to the products, the customers who have profound insight seek the overall effectiveness of maximizing products purchasing and brands spreading, and the dominant sellers can maximize the two-stage total profits through using reasonable pricing strategies. Huang [3] and Krishnan et al. [4] elaborated the pricing strategy and analyzed four strategies, such as new product pricing strategy and discount strategy, which can help enterprises make full use of pricing strategy to provide reasonable market prices to obtain the most competence. But all the above literatures just study the traditional price strategy and cannot explain the popular free phenomenon nowadays. 1P business model proposed by Wang [12, 13] breaks the pricing space of the business strategy by integrating the third party to obtain a new competitive advantage, with putting the pricing into the center position and innovatively stressing positioning and locking the third party to pay, which can make profit with the price that is lower than the average cost and can promote with the price that is higher than the expectation of customers.

In this paper, through the deep research on price discrimination theory and crowdfunding pricing mechanism, we not only consider the influence of crowdfunding price on needs, but also further study the effect of network brand spreading and advertising, constructing the maximum profit function of the product crowdfunding in the two occasions to obtain the pricing expression under the constraint for financing amount. This paper also studies the pricing theory of business model proposed by Wang [13] and applies it into the crowdfunding business model with drawing the corresponding pricing strategy.

3 Crowdfunding Pricing Model

3.1 Crowdfunding Pricing Model for Price Discrimination

According to the level of discrimination, the price discrimination can be divided into three levels. The first-degree price discrimination, also known as complete price discrimination, regulates the price according to the possible maximum amount of money paid by each buyer to obtain the full consumer surplus. Based on different consumptions or "segments" to obtain a different price, the second-degree price discrimination only regulates different prices for different segments and thus obtains part of the consumer surplus. The third-degree price discrimination classifies the customers according to a particular standard to make clear which category each customer belongs to, with taking different prices in different markets or to different customer groups for one kind of product.

The first and the second degrees of price discrimination enable the manufacturers to grab the consumer surplus wholly or partly, and this part of the consumer surplus will be transferred into profits, achieving the principle of effective allocation of resources $P = MC$ (equilibrium price equals to marginal cost. That is the requirement of maximizing profits). However, the first-degree price discrimination cannot be fully realized in reality. The pricing of crowdfunding products will usually take the pricing strategies of the first and second degrees of price discrimination, using direct information on needs to choose the consumers indirectly through the choices done by them about different packages. The inverse elasticity rule is taken for different consumer groups, meaning that the consumers who have a high price elasticity will be charged low and the consumers who have a low price elasticity will be charged high. The maximization of profits can be obtained through the price combinations.

By the function of demand and price

$$Q_d = d(P).$$

The demand function of the consumers belonging to different levels in the price combinations is

$$Q_n = (A_n - P_n)/E_p,$$

here A_n is the price that the consumers are willing to pay and E_p refers to the price elasticity of demand.

Assuming that β is the variable cost component of producing a unit of this product and the initial funding amount K is the fixed cost amount put by the firm to produce the required product smoothly. The total cost of product is

$$C = K + \beta \sum_{i=1}^{n} Q_i.$$

The total profit of the project producer is

$$R = \sum_{i=1}^{n} P_i Q_i - K - \beta \sum_{i=1}^{n} Q_i.$$

Only if the total payment of the participant is not less than the total amount of financing K can the raisers get the money from the platforms smoothly. It requires:

$$\sum_{i=1}^{n} P_i Q_i \geq K.$$

To maximize the profits of crowdfunding, it is required to consider how to design the relations between various levels of prices and quantities, that is, to figure out the maximum profit objective function with some constraints

$$\max\ R = \sum_{i=1}^{n} P_i Q_i - K - \beta \sum_{i=1}^{n} Q_i$$

$$\text{s. t.}\ \begin{cases} P_n = A_n - E_p Q_n, \\ \sum\limits_{i=1}^{n} P_i Q_i \geq K, \\ P_i \geq 0. \end{cases}$$

3.2 The Crowdfunding Price Model Concerning with the Advertising Effect

As a result of extensive and effective brand and spreading effects, the product crowdfunding not only can sell the products but also reach the goal of advertising through the release on the Internet platform. Compared with the common advertising tunnels, the influences of crowdfunding include a lower price and sound effects. However, considering that the effect is related much to the demand and price, it is still necessary to research the pricing strategy prudently to maximize the profits.

Assuming the same amount of customer traffic, denote the conventional unit of product advertising budget by G_r and the cost of crowdfunding platform by K_c, the function relation between effect profits and the demands of the rather common advertisements is

$$G_c = G_r \sum_{i=1}^{n} Q_i - K,$$

where the function relation between the price and the demand is

$$Q_n = (A_n - P_n)/E_p.$$

A_n refers to the price which the crowdfunding consumers are willing to pay and E_P is the price elasticity of demands.

Assuming that β is the variable cost component of producing a unit of this product and the initial funding amount K is the fixed cost amount that the firm can produce the required product smoothly. The total cost of product is

$$C = K + K_c + \beta \sum_{i=1}^{n} Q_i.$$

The total profit of the project producer is

$$R = \sum_{i=1}^{n} P_i Q_i - K + G_c - \beta \sum_{i=1}^{n} Q_i.$$

Only if the total payment of the participant is not less than the total amount of financing K can the raiser get the money from the platform to produce smoothly. It needs to meet with:

$$\sum_{i=1}^{n} P_i Q_i \geq K.$$

Taking the advertising effect into account, it is needed to figure out the maximum of profit objective function with constraints to obtain the maximum profits of the crowdfunding

$$\max R = \sum_{i=1}^{n} P_i Q_i - K - K_c + (G_r - \beta) \sum_{i=1}^{n} Q_i.$$

3.3 The Crowdfunding Pricing Model Using the 1P Theory

1P theory refers that the enterprises can expand the profit space through integrating the third party to benefit mutually, to participate in income payments and cost sharing and to break the upper and lower limits of strategic pricing. The third party represents the value of all associated networks, and is the value of the associated network of the enterprise and its customers and partners.

Crowdfunding as a business pattern can also be explained by the 1P theory which wields the third party to pay, and all the crowdfunding patterns can be categorized in the 1P theoretical business model. The product crowdfunding views that the investor as a third party to share some or all of its fixed cost, channels and promotional cost, thereby reducing the cost of the product and improving the total profit of the product. Therefore, the product is also the value logic and trading structure in which the companies, customers and other third-party partners can jointly create and trade value to make mutual benefits.

In the 1P theory, the 4Ps are separated according to the nature of the income and cost, where the price P is the income per unit of product; the product P, the channel P and the sales promotion P are referred to as 3P, which is the unit cost (that is, the average cost AC and $3P = AC$). π refers to the unit profit, and Q means the production, and $\pi = P - 3P = P - AC$, therefore the total profit is $\pi = Q(P - AC)$. Here, $P - AC$ is both the pricing space and profit margin.

According to the linear value hypothesis of strategic marketing, the product price P is paid by the target customer C, so $P = PC$. The unit product cost AC is burdened by the enterprise E itself, so $AC = ACE$, and the product price must be greater than or equal to the average cost $P \geq AC$, that is $PC \geq ACE$, or there is money loss. When the profit is zero, $\pi = P - AC = 0$, i.e. $\pi = PC - ACE = 0$, $PC = ACE$.

The introduction of the third-party investors in the product pool will make the unit product price P no longer equals to the target customer's payment price PC, for it should be added with the price of PB paid by the third-party investors B, that is, $P = PC + PB$, $PC = P - PB$; when the profit is zero, $\pi = (PC + PB) - AC = 0$, $PC = AC - PB < AC$. This proves that even if the price of the product sold to the target customer is less than the price of the product and the average cost, it is also profitable, and how much the profit will be depends on the price paid by the third party and the number of the shared cost.

The theoretical business model of 1P whose pricing space breaks the upper and lower limits of pricing is shown in Fig. 1.

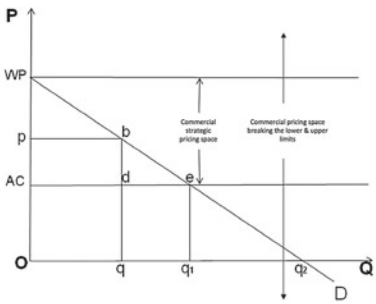


Fig. 1. The business pricing space model of 1P theory

4 Crowdfunding Pricing Strategies

4.1 The Crowdfunding Pricing Strategy Using the Price Discrimination

The crowdfunding products are faced up with different consumer groups who have different demand elasticity to products, creating some conditions for the implementation of price discrimination. The consumers who have higher elasticity can be charged with lower prices to increase sales to gain profits while the consumers who have lower elasticity can be charged with higher prices to raise the profit without decreasing the sales obviously. Considering from elasticity, the supporters of crowdfunding are willing to pay higher prices. When considering uncertainty about the quality of products, supporters of the crowdfunding will face a greater degree of uncertainty, then they need to be given with discounts and the price should be lower. Hence, it is not clear which part of the consumers will face higher prices in the crowdfunding projects. In addition, because the consumers in the crowdfunding projects can choose whether to participate in at their own accordance and the products are marked with personalized features, it is assumed that consumer groups do not trade with each other, and producers can implement price discrimination.

Obviously, price discrimination can make the sellers benefit as much as possible, for the consumer surplus belonging to the buyers will also be transferred to the sellers through price discrimination. Price discrimination is economically efficient. The maximum benefit of sellers equals to the value of social welfare maximization through price discrimination. Of course, the sellers have to be able to distinguish different characteristics of the buyers to make price discrimination work. This difference may exist in the buyers' demand intensity, the quantity of purchase or the price elasticity of demand. The basic principle of implementing price discrimination is that the marginal returns in different markets are equal and equal to marginal cost. The crowdfunding management can regulate a higher price for the market with less price elasticity of demand, implementing the strategy of "less sales but more profit", while a lower price can be regulated for the market with greater price elasticity of demand, and the maximum profit can be obtained through the implementation of "less profit but more sales".

4.2 The Crowdfunding Pricing Strategy Considering the Advertising Effects

Plenty of crowdfunding products are marked as innovative products. If the products are new to the market, it can be considered in the price combination about the three pricing strategies including skim-milk pricing, penetrating pricing and satisfying pricing. Also, as a new product to come out, to select the crowdfunding is also a working advertising way which can lead to a sound effect. Thus, the advertising effect needs to be considered in the crowdfunding pricing strategy, and the ways mentioned above should be implemented flexibly according to the actual situation to achieve profit maximization.

Skimming pricing is a high-priced pricing strategy in the early stage of a new product, provided that the product's quality and image must support the high price of it. Also, consumers have a strong desire for consumption and there is no potential competitor in the market. In this case, consumers are willing to pay for high prices; also, the elasticity of demand, sale and the overall profit income are rather high.

Penetrating pricing strategy is a low-priced strategy when a new product first comes into the market, in which many consumers will be attracted by the low price to win a sound reputation and a comparatively large market. The prerequisite of this pricing is: the market is highly price-sensitive and the cost of producing and selling will be reduced with increasing sales. Also, the competitors can be expelled through a low price. Such a disruptive and innovative pricing method tends to stimulate consumers' willingness to buy and bring about sound advertising effects, resulting in more profits.

Satisfying pricing strategy is a strategy between skimming pricing and penetrating pricing. It is an intermediate pricing which can satisfy both producers and consumers. The prerequisite of such a pricing is that this product is comparatively mature and the consumers are familiar with the product. Also, the information is symmetry. The consumers can buy this product according to their own preferences. If the consumers are satisfied with the price, more attention of them will be placed on the product experience, then the high-quality product will win a good name, and the manufacturers will thus obtain rich profits.

In short, it needs to take the actual situation of crowdfunding products into account to choose skimming pricing, penetrating pricing or satisfying pricing. Considering the cost of advertising and error-trying, the maximum product profit can be gained with using a crowdfunding way to sell and make a sound pricing strategy because of the novelty and advertising needs of these products.

4.3 The Crowdfunding Pricing Strategy Using the 1P Theory

The 1P theory claims that any business model is a revenue-cost transaction structure in which customers, third-party associated customers, and third-party production partners participate in co-creating and trading network values. In the business model of the product crowdfunding, the consumer may be a producer or the producer may be a consumer at the same time. The third-party associated customers who pay the prices can be either consumers or enterprises; the third-party production partners who share costs can be either a business or a consumer. Transaction cost is the most common network value in the network economy, and the emergence of the third parties makes transaction cost lower. The reduction in the number of transaction equals to the result of the number of brands multiplied by the number of customers minus the number of brands then adds the number of customers, and the difference will be multiplied by the cost of each transaction. The result is the transaction cost saved by the third-party integrated brands and the customers. Therefore, the number of the third parties can be expanded through the pricing strategy to lower the transaction costs in a greater degree to maximize profits.

The business model innovation of crowdfunding should break the lower and upper limits of the commercial competitive pricing to access to new competitive advantages and expand profit space through discovering, creating, trading network values and integrating the third-parties' paying. In the formulation of pricing strategies, it should be focused on discovering, creating and capturing the network values, paying attention to the cooperation among businesses, customers and partners to jointly benefit from expanding the total values. Also, it needs to be considered to use the ecological network thinking to re-examine the economic nature of the businesses, customers, competitors, suppliers and even the whole commercial environment to assess the values of the associated network within or outside the enterprise and find out the high-volume crowdfunding-targeted customers to jointly create and trade network values, achieving the purpose of win-win cooperation.

5 Conclusion

5.1 Theoretical Contribution

In this paper, we studied profoundly the price discrimination theory and the pricing mechanism of the crowdfunding pricing. By considering the effect of the crowdfunding pricing on demand to compose the profit function and the constraint conditions of the financing, it was figured out the crowdfunding pricing expression concerning with the price discrimination. Compared with the traditional sales modes, the total sales of the crowdfunding products remain the same, while the proportions of crowdfunding consumers and conventional consumers change oppositely. The crowdfunding mode can distinguish the different preferences of different consumer groups, according to which the market's attitude towards the products is likely to be judged. The unit price given by crowdfunding consumers is higher than that of conventional consumers. The crowdfunding raisers are faced up with a funding ceiling, and if the ceiling is touched, the crowdfunding will be a failure. Not only can the new medium and small enterprises finance without much difficulty by crowdfunding, but also more profits can be achieved for the businesses compared with the traditional sale modes.

Also, we further studied the maximum profit function of the product crowdfunding with the advertising effects. Finally, we studied the pricing theory of the business model in which the third parties pay in the theory of 1P, and then applied it to the business model of crowdfunding to get the corresponding crowdfunding pricing strategies, aiming to obtain the maximum profit. The paper is expected to contribute to the theoretical research in the crowdfunding financing and pricing field, and also provide references for enterprises and entrepreneurs to make decisions on pricing strategies for their products.

5.2 Marketing Significance

Pricing strategy is an important part of the marketing combination for a business. Enterprises need to prudently formulate the pricing strategies of the products when making the crowdfunding plans. Only by innovating and conforming to the objective laws of the market economy can the enterprise win more profits in the target market and achieve a better reputation and advertising effects. Making the right crowdfunding pricing strategy is not only an important factor in maintaining the interests of manufacturers and customers, but also the key to overcome competitors, develop and consolidate the market.

5.3 Limitations and the Future Research Prospects

Due to the short rising period of the Internet crowdfunding, the domestic and foreign literature about its product pricing is correspondingly scarce. Crowdfunding product pricing research also belongs to the emerging field, which has great research value and profound potential. Moreover, with the rapid iteration of technologies and business models, the industry of crowdfunding is changing incessantly and the pricing strategy research should follow the development of the situation. This paper gives the theoretical expression of the crowdfunding pricing using the price discrimination as well as the maximum profit function considering the effects of advertising, and there may be more factors which should be taken into account when planning the actual cases.

With the support of the policies, the developing potential of crowdfunding model is not to be limited, especially in the items of culture innovation and technologies. Young people are more likely to express themselves and communicate on the Internet, thus, the consumer groups of the products of culture innovation and technologies match well with the main active users on the Internet. With the help of the crowdfunding platforms and the social network, the crowdfunding projects can be focused as soon as possible. Not only can the money be obtained but also the sound advertising effects can be achieved. If we introduce the analyzing technologies of large data into pricing strategies, it may be more accurate in terms of dimensions and data processing to make the strategies more effective.

Acknowledgements. This project was supported by the National Natural Science Foundation of China (11501035, 70831005) and the Key Project of China Petroleum and Chemical Corporation (GJ-73-0706).

References

1. Bayus BL (1997) Speed-to-market and new product performance trade-offs. J Prod Innov Manag 14(6):485–497
2. Hu M, Li X, Shi M (2014) Product and pricing decisions in crowdfunding. Soc Sci Electron Publishing 34(3):331–345
3. Huang ZY (2012) Discussion on product pricing strategies. J Nanchang Coll Educ 27(11):12–18 (in Chinese)
4. Krishnan TV, Bass FM, Jain DC (1999) Optimal pricing strategy for new products. Manag Sci 45(12):1650–1663
5. Lawton K, Marom D (2013) The crowdfunding revolution: how to raise venture capital using social media (ebook). McGraw Hill, New York
6. Luo Y, Gao L, Huang J (2015) Price and inventory competition in oligopoly tv white space markets. IEEE J Sel Areas Commun 33(5):1002–1013
7. Nocke V, Peitz M, Rosar F (2011) Advance-purchase discounts as a price discrimination device. J Econ Theor 146(1):141–162
8. Pigou AC, Todd AJ (1934) The economic of welfare. Am J Sociol 19(1):23–30
9. Ping YY (2015) The pricing mechanism of commodity crowdfunding. Zhongnan Univ Econ Law 40(3):23–30 (in Chinese)
10. Rochet JC, Thanassoulis JE (2016) Stochastic delivery and intertemporal price discrimination with multiple products. Social Science Electronic Publishing
11. Sahm M, Belleflamme P, Lambert T et al (2014) Corrigendum to “crowdfunding: tapping the right crowd”. J Bus Ventur 29(5):610–611
12. Wang JG (2015) An investigation into modeling marketing. J Peking Univ 52(4):95–110 (in Chinese)
13. Wang JG (2016) The classification, innovation and design of business model in business ecosystem. Peking Univ 6(1):1–17 (in Chinese)
14. Zhang MX, Lei M, Zheng XN (2013) The impact of online word-of-mouth dissemination on the oligarch seller’s pricing strategy. Mark Sci 9(2):71–89 (in Chinese)

The Support System of Innovation-Driven Strategy in Private Enterprises: A Theoretical Model

Tao Xie[1], Guichuan Zhou[1](✉), and Shuangyi Zheng[2]

[1] Business School, Sichuan University, Chengdu 610065, People's Republic of China
zhouguichuan@vip.163.com

[2] School of Management of South-Central, University for Nationalities, Wuhan 430074, People's Republic of China

Abstract. The theory of Innovation System has been used to understand the complexity and instability of innovation. Recently Strategy management researchers have found Innovation-Driven Strategy is transferred from national level to enterprise level. The dualistic Innovation System is dominated by government agencies and enterprises in parallel, which is less conducive to innovation of private enterprises. The dualistic main-body Innovation System is converted into the multi-nested Innovation Support System by building a Support System of Innovation-Driven Strategy (SSIDS) in private enterprises. Through analyzing the connotation, structure and characteristic, a new theoretical model is pioneered that serves as a guide for innovation strategy in private enterprises.

Keywords: Private enterprises · Innovation-driven strategy · Support system

1 Introduction

Managers should put additional importance on innovation that plays a significant role as a bridge between competitive strategies and firm performance in a developing economy environment [1]. As stepping into the "new normal", the development of economy is not investment-driven, but innovation-driven in China, which reflects the reconstruction of economic engine and switching of developmental strategy [12]. Much of research has examined that private enterprises (PE) are superior to state-owned enterprises (SOE) about innovative ability [8]. According to the statistics, private enterprises have more R&D projects, R&D investment and patents than state-owned enterprises. Therefore, private enterprises with innovative potential play a leading role in innovation-driven strategy. However, market mechanism with inherent weaknesses has blocked the development of innovation in China. "Perfect market" also breeds Opportunism which jeopardizes innovative performance of PEs [4]. Although with huge innovative potential, PE faces systemic risk. When in the immature market, entrepreneurs improve transformational leadership with the expectation of favorable

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_67

business environment that the government creates. The dualistic Innovation System dominated by government agencies and enterprises in parallel is extensively implemented in China [16]. There are some divergences concerning relationship between the government and enterprises, which roots in complexity and instability of innovation. The theory of Innovation System cannot solve new problems of PEs.

How does PE build the internal support system to keep innovative performance? How does the government improve the business ecosystem without interference to PE This article will solve these problems and develops a theoretical model of Innovation-driven Support System, which can change traditional innovation paradigm and offer a framework to further empirical studies. To do this we build our discussion on extensive relevant literatures and enterprise practices. Using methods and theories of the system, the Support System of Innovation-Driven Strategy (SSIDS) is built.

In the following sections, we first review current researches and state deficiencies. We then go on to analyze connotation, structure and characteristic of SSIDS. Finally, we discuss the implications of our theoretical model for managers and provide directions for further research.

2 Literature Review

The inflexion point theory demonstrates that according to the decreasing law of marginal utility, contribution degree of basic elements, such as natural resources and capital, present descending trend. Therefore, the long-term development of economy depends on intellectual and technological factors. The inflexion point theory promotes innovation-driven theory, which manifests investment-driven strategy is replaced by innovation-driven strategy. The concept of innovation-driven was initially used by Porter [13]. Porter argued that economic development could be divided into four periods: factor-driven stage, investment-driven stage, and innovation-driven, wealth-driven stage. In innovation-driven stage, the effect of basic factors and investment is led by innovation and a system of innovation-driven strategy taking technological innovation as core is built.

After that innovation-driven development became national strategy, Chinese academics conduct quiet a few examinations regarding innovation-driven strategy. Liu [11] contended that the innovation-driven development means the change of economic engine in China. And he further demonstrated that the growth of economy has transformed from depending on learning and imitation into relying on self-designing, R&D, knowledge-creating. In addition, Zhang [17] analyzed the main features of innovation-driven development, i.e. people oriented, acquiring first-mover advantage and entrepreneur-driven. Li and Wang [10] proposed internal innovation-driven elements including enterprise culture, leaders full of entrepreneurship, expected return; and external innovation-driven elements including technical progress and institutional environment. Meanwhile, WANG Tao and Qiu [16] emphasized that the innovative environment dominated by entrepreneurs is immature and constructed innovation-driven strategy of "two-way drive" effect model that dominated by government and enterprises in parallel.

Innovation origins in Economics, gradually progresses to Management. Meanwhile, innovation-driven development transfers from national strategy to enterprise strategy and entrepreneurs play a major role in innovation. Implementing the innovation-driven strategy is complicated and nonlinear systems engineering. Therefore, many studies have advanced the combination between innovation and system and developed the theory of Innovation System, such as National Innovation System (NIS) [15], Regional Innovation System (RIS) [5], Technological Innovation System (TIS) and Innovation Ecosystem (IE) [2,6]. However, more recent studies argued that Innovation Support System (ISS) is different Innovation System and suggested the performance of innovation actors is restricted by environment, in China. He and Li [7] proposed how the Innovation Support System is established. In the process of innovation, various innovative elements generate and spread new knowledge and technology by interaction to form a network system.

In summary, innovation-driven strategy not only depend on the core competency of enterprises, but need a favorable business environment. Previous studies built the Innovation System that is a dualistic model, which cannot clarify the role of the government and enterprises. The Innovation Support System provides a feasible theoretical framework for innovation-driven strategy in private enterprises. It is how private enterprises effetely combine their core competency and environmental support from the government mostly, which is the deficiency of present literature.

3 The Structure of SSIDS Model

It is complex system engineering for private enterprises to implement the innovation-driven strategy, which is influenced by various factors inside and outside the system. Elements form inherent complex relationships by interaction, such as resources, capabilities and environment, which is regarded as a system. Therefore, it is necessary to build a system to support the innovation-driven strategy in private enterprises.

3.1 Strategic Support System

Strategic support system is a collection of subsystems composed by a set of key managerial issues, which support the realization of strategic objectives. This system can be divided into internal and external support subsystems. Internal strategic support subsystem is constituted by internal key managerial issues within the process of strategic implementation and usually includes three parts: impetus subsystem, behavior subsystem and capability subsystem. Specifically, the focus of impetus subsystem is to form power of strategic implementation; the focuses of behavior subsystem include deciding, organizing, motivating and controlling behaviors during strategic implementation; the focuses of capability subsystem is mainly about the optimization of enterprise capabilities including strategic management, R&D, capital operation and marketing. External strategic support subsystem is the collection composed of various external factors and their interactions in the process of implementation of strategies. It mainly refers to external supportive environment, for example, market competition environment, resource environment, policy & laws environment and public service environment.

Support System of Innovation-Driven Strategy (SSIDS) in private enterprises refers to the organic integration and coupling operation of enterprises' internal and external supporting elements. In this system, various elements promote and restrict mutually to seek the balance point supporting the implementation of private enterprises' innovation-driven strategy. Therefore, this system will help to enhance the willingness of private enterprises to implement innovation-driven strategy, facilitate the exchange of internal and external factors such as material, energy and information during strategy implementation process, and further improve their innovative performance.

SSIDS in private enterprises includes two basic parts: internal and external support system. It covers every supportive element inside and outside enterprises during the process of the implementation of innovation-driven strategy, centering on the improvement of innovative performance.

3.2 Internal Support System of SSIDS

Internal support system is a relatively stable network system formed endogenous in private enterprises, consisting of innovation support elements and their mutual relationships. The process of innovation-driven development in private enterprise can be described as: generate innovation impetus, implement innovation behavior, form innovation capability, which leads to the formation of innovation impetus subsystem, innovation behavior subsystem and innovation capability subsystem, as illustrated in Fig. 1.

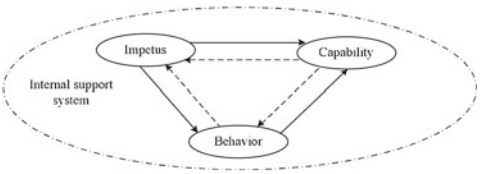


Fig. 1. Internal support system

(1) Impetus Subsystem

The degree of innovation impetus directly influences the speed and scale of innovation in private enterprise. According to the different sources, innovation impetus can be divided into internal impetus supply and external impetus supply. As showed in Fig. 2, internal impetus supply includes resource, entrepreneurship and corporate culture; external impetus supply includes technological advancements, market demands, competitive pressure and the support from government. Impetus of resource comes from continuously innovation resources' injection to internal innovation system by enterprises. These innovation resources include technical personnel, innovative capital and technical equipment etc. Abundant resource can improve the success rate of innovation, increase innovative returns, and thereby strengthen the willing of implementing innovation strategy.

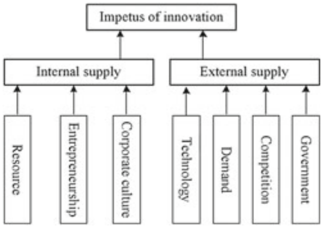


Fig. 2. Impetus subsystem

Entrepreneurial spirit becomes the impetus of continuous innovation in private enterprises through entrepreneurs' self-driven and self-motivation. The innovative spirit of entrepreneurs can motivate more members to participate the practice of innovation-driven strategy and further improve the market performance.

Enterprise innovative culture can influence staff's values, thoughts and behaviors, which infect and motivate innovation of organization, intrinsically and intangibly. Innovative culture performs its function as guiding, constraint, coherence and motivation [14].

When converted to differentiated products or services, achievements in scientific and technological progress can bring higher monopoly profits. Also that motivates enterprises consistently to absorb new results of scientific and technological progress and to re-innovate. The impetus and influence of technological progress performs as: the induction of new technology ideas, technical track and technical expectations.

Market demand is the basic starting point of implementing innovation-driven strategy for private enterprises, which provide new market opportunities and innovative ideas for enterprises. In fact, the market demand has the function of guiding the direction of technological invention, testing the results of technological innovation and providing technical promotion space.

The pressures of competition enhance the crisis awareness of private enterprises. Meanwhile, the incentive of innovation can also be triggered by more market share and monopoly profit. Under fierce competition, the private enterprise will have more intense willing of innovation, more emphasis on market intelligence, more advanced concept and shorter product update cycle.

The government support can make up inherent weaknesses of market. Innovation activities characterized of high risks and externalities, lead to the decline of innovative impetus of private enterprises. To reduce this negative effect, the government should give private enterprises appropriate support such as making up for the overflowed loss of innovation, building innovative infrastructure and sharing risk in innovation.

(2) Behavior Subsystem

When facing various stimulations from external environment, private enterprises make responses by implementing innovation-driven strategy. The behaviors of innovation-driven strategy integrate the interaction of macro and micro, vertical and horizontal, external environment and internal factors, and finally, form a closely related dynamic open system. Figure 3 illustrates the composition of innovation behavior subsystem of private enterprises, the subsystem includes four aspects: policy-making, organization, motivation and control. According to the theory of organizational behavior, the evolution direction of private enterprises' innovation behavior is: policy-making behavior organization behavior motivation behavior, moreover, the control behavior is penetrating in each behavior stage. As the elements of the system, different behaviors also have reverse exchange of information and energy. The detailed path is shown in Fig. 3.

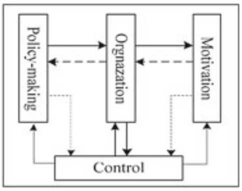


Fig. 3. Behavior subsystem

Behaviors of policy-making imply that managers revise innovation utility function and the risk function by collecting information and decide whether to carry out the innovative activities, which intends to reduce the uncertainty during innovation process. Managers with strong policy-making ability can not only grasp the market information accurately and enhance the willingness of implementing innovative activities, but also can form effective methods and means to improve the success rate of innovative activities.

Behaviors of organization refer to that private enterprises conduct optimized recombination of innovation elements such as technical knowledge, innovation capital and intelligent team in the process of strategic implementation [3]. The quantity, quality and synergy of these factors determine the output and risk of innovative activities. Technical knowledge is the foundation of innovation, which determines the feasibility and risk of innovation. Capital is the blood of innovation, which ensures stability and continuity of the innovative activities. Intelligent team is the brain of innovation; high quality and compound innovative team provides innovation activities the insurance of science and sustainability.

Behaviors of motivation include a series of specific regulations and reward-punish system, by which behaviors of staff are in line with the goal of innovative activities, and their potential are exerted fully. The implementation of innovation-driven strategy is catalyzed by innovation incentives such as spiritual incentives and material incentives. The intensity of incentive will affect the desire and responsibility of innovative activities. Meanwhile, the way of incentive will influence the relationship between innovation behavior and expected returns.

Behaviors of control imply that managers diagnose, adjust and feed back all other behaviors, which can reduce the loss of error behavior, and improve the incomes of correct behavior. Specifically, behaviors of control includes performance evaluation, risk monitoring, modification, feedback and so on. Remarkably, control behavior should run through the whole innovation process because innovation activities are characterized of extreme complexity and uncertainty.

(3) Capability Subsystem

Innovation capability refers to the probability, efficiency and continuity that enterprises carry out innovative activities and obtain innovative income. As shown in Fig. 4, the capability is mainly embodied in knowledge stock and innovative output. With strong innovation capability, enterprise can constantly develop new products, expand the development of the enterprise space, and improve the level of its profitability to maintain strong innovative motivation [9].

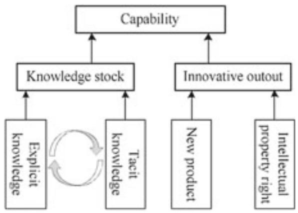


Fig. 4. Capability subsystem

Knowledge stock consisting of explicit knowledge and tacit knowledge is the key element of enterprise growth, resource allocation and innovation. The flow and upgrading of knowledge can nurture flexible thoughts and enhance the recognition of innovation strategy. It is the basis of implementing innovation-driven strategy to improve quantity, quality and structure of the knowledge stock.

Innovative outputs, generally speaking, include new products and intellectual property rights. Once the innovation results are converted into new products with competitiveness, enterprises can attain revenue from innovation. The higher the market value of new products is, the stronger the willingness of enterprise to innovate will become. As the means to ensuring the benefits of innovation, intellectual property rights have an important impact on the sustainability and stability of the strategy. The more intellectual property rights you have, the more security your innovative performance will gain.

3.3 External Support System of SSIDS

There are lots of external elements of enterprise that are closely related to innovation-driven strategy. External support system is exactly consisted of these elements and their mutual relationship together. It is necessary to implement the innovation-driven strategy and can be generalized as the environmental subsystem of innovation. As shown in Fig. 5, the system commonly includes market environment, policy environment and service environment.

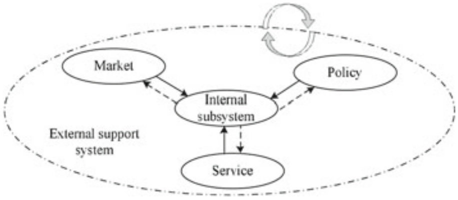


Fig. 5. External support system

(1) Market Environment

Influenced by historical reasons and institutional factors, Chinese private enterprises are facing more and special market environment. Market environment mainly includes the market rule, the fairness of market competition and the degree of government's intervention to market. With the help of adequate market rules, essential market ethics and credit, less intervention from government, private enterprises can compete in a fair market environment and further improve their willingness and performance of innovation.

(2) Legal Environment

Laws & regulations play an important role of ensuring and adjusting innovative activities of enterprises. Therefore, government should focus on creating suitable environment and help to promote innovative activities in enterprise. For example, the authority can exert its influence on launching, organizing and implementing of enterprises' innovative programs by constantly adjusting and perfecting the laws and regulations on intellectual property rights and the standards of environment & energy. These laws, especially intellectual property protection regulations, can help private enterprises to enhance technical barriers, protect the property rights of innovation, and improve performance from innovation.

(3) Service Environment

The insurance of technology, personnel and capital is necessary for private enterprises to implement innovation-driven strategy. Therefore, service environment is vital for innovative activities. First, it can serve as bridge connecting technology and economy by providing private enterprises mature intermediary institutions technology and market. Second, it helps to introduce and select all

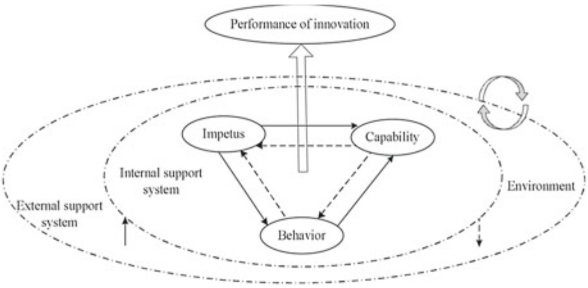


Fig. 6. Support system of innovation-driven system

kinds of innovative talents for private enterprises. Finally, it promotes the formation of a sound financial market so that private enterprises can raise innovation funds from this important capital source.

Based on above analysis of internal and external support system, we build a new support system of innovation-driven strategy in private enterprises with the purpose of improving performance of this strategy (as illustrated in Fig. 6). The support system of innovation-driven strategy totally differs from the traditional innovation systems that define the role of the government and private enterprise. The SSIDS is an open, dynamics and complex system as it involves a wide range.

4 Conclusion

This article analyzes internal support system form innovative impetus, behavior and capability, discusses external support system from market mechanism, laws and regulations, service. Based on internal support system and external support system, a new theoretical model, SSIDS, is built, which is conductive to innovative practices of private enterprises.

Innovation is unpredictable system engineering. According to Schumpeter's theory of innovation, entrepreneurs are situated in primary status. However, at the initial stage of the market, Chinese private enterprises lack of mature entrepreneurs. Therefore, traditional theory of Innovation System insists that both the government and enterprises are situated in primary status, which is harmful to development of Chinese entrepreneurs. Contrary to the dualistic main-body Innovation System, the SSDIS with a multi-nested structure is instrumental and assistant. In the SSDIS, the government provides service and supervision; private enterprises carry out innovative practice. This article builds a theoretical model and change innovation paradigm. Nevertheless, many empirical studies are necessary to the validation and optimization of SSIDS model. How the SSIDS model operates is a valuable question for further research.

Acknowledgements. We thank the support of the Projects of the National Social Science Foundation of China (16AGL003), Social Science Planning Annual Project of Sichuan Province (SC16C011), Soft science research plan project of Sichuan province

(2017ZR0098, 2015ZR0091), Soft science research project of Chengdu city (2015-RK00-00156-ZF), The Central University Basic Scientific Research Business Expenses Project of Sichuan University (skqy201623, skqy201209), Projects of the Center for systems science and enterprise development (xq15B11).

References

1. Bayraktar CA, Hancerliogullari G et al (2016) Competitive strategies, innovation, and firm performance: an empirical study in a developing economy environment. Technol Anal Strateg Manage
2. Carlsson B (1991) On the nature, function and composition of technological systems. J Evol Econ 1(2):93–118
3. Chakraborty S, Thompson JC, Yehoue EB (2016) The culture of entrepreneurship. J Econ Theory 163(3):288–317
4. Chen B (2014) On the connotation and characteristics of innovation driven and the condition of realizing innovation driven: from the perspective of realizing "China dream". J Theor Ref 7:123–134 (in Chinese)
5. Cooke PN, Heidenreich M, Braczyk HJ (2004) Regional innovation systems: The role of governances in a globalized world. Eur Urban Reg Stud 6(2):187–188
6. Edquist C (1997) Systems of innovation: technologies, institutions and organizations. Soc Sci Electron Publ 41(1):135–146
7. He M, Li F (2009) Innovation supporting system of Zhongguancun science park. J Beijing Inst Technol (Soc Sci Ed) 2:46–55 (in Chinese)
8. Jiang S, Gong L, Wei J (2011) The path to the catch-up of the innovative ability of the late comers of enterprises in the transitional economy: comparing SOE with POE. Manage World 12:43–72 (in Chinese)
9. Lancker JV, Mondelaers K et al (2016) The organizational innovation system: a systemic framework for radical innovation at the organizational level. Technovation 52–53:40–50
10. Li Y, Wang J (2014) The research on the relationship between driving factors and performance of innovation in the context of vertical integrated industry chain. Res Finan Econ Issues 7:21–29 (in Chinese)
11. Liu Z (2011) From late-development advantage to first-mover advantage: theoretic thoughts on implementing the innovation-driven strategy. Ind Econ Res 4:65–72 (in Chinese)
12. Liu Z (2015) Innovation-driven strategy in new high-level opening-up. J Nanjing Univ (Philos Humanit Soc Sci) 2:37–47 (in Chinese)
13. Michael P (1990) The competitive advantage of nations. Ashgate, Brookfield
14. Mir M, Casadesüs M, Petnji LH (2016) The impact of standardized innovation management systems on innovation capability and business performance: an empirical study. J Eng Technol Manage 41:26–44
15. Trevor M (1989) Technology policy and economic performance. lessons from Japan. R&D Manage 19(3):278–279
16. Wang T, Qiu G (2014) A research on "bidirectional-driven" effects of innovation-driven strategy. Technoeconomics Manage Res 6:44–57 (in Chinese)
17. Zhang L (2013) Innovation-driven development. China Soft Sci 1:80–88 (in Chinese)

The Emission Reduction and Recycling of Coal Mine Water Under Emission Allowance Allocation of Government

Ning Ma and Shuangyi Zheng(✉)

Business School, Sichuan University,
Chengdu 610065, People's Republic of China
zhengshuangyi@scu.edu.cn

Abstract. Water shortage and water pollution problems have become increasingly severe, especially in these areas with the coal mining industry. In this paper, a bi-level multi-objective optimization model based on the trade-and-cap mechanism with fuzzy coefficients is developed to deal with environmental water problems in large scale coal fields, in which both the emission right and mine water recycling are considered. This model fully considers the relationship between the government and the coal mines and also the equilibrium between economic development and environmental protection. To deal with the inherent uncertainties, the model is defuzzified using a possibility measure, and a solution approach based on the Karush-Kuhn-Tucker condition and interactive fuzzy programming is designed to search for the solutions. A case study is presented to demonstrate the practicality and efficiency of the model. The results showed practicality and efficiency in the proposed model of reducing the coal-water conflict. Finally, a comprehensive discussion is provided and some propositions is given as a foundation for the proposed management recommendations.

Keywords: Water environment · Coal mining · Mine water recycling · bi-level multi-objective programming

1 Introduction

Coal is still the most abundant fossil fuel in 2015 and it provided around 30.0% of the global primary energy need, though oil and natural gas reserves have increased over time. As such an important fossil fuel contributing to worldwide energy generation, coal will still takes a huge market share in world energy and coal mining is still a major industry for developing economy. But coal mining can cause a lot of pollution for water, soil and air environment, and these ecological environment issues will deepen the contradiction between the land and humans, and constrains economic and social development. However, for safety, significant quantities of groundwater are discharged in underground mining [17,19]. If

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_68

untreated, the water combining with heavy metal ions such as iron and mercury ions would definitely pollute the groundwater [7], which has many pollutants and waste the cherish water resource [16,18]. Thus, it is of importance for the coal mine to control the mine drainage.

In order to mitigate the environmental damage, several policy instruments have been developed to attempt to reduce pollution emissions, such as emission taxes, command-and-control, and cap-and-trade [5,6,9]. The cap-and-trade mechanism, also known as the emission trading scheme (ETS), is an application of Coase [3], and has showed its effectiveness of controlling emissions and has been successfully applied in practice [2]. Therefore, this paper adopts this mechanism to control the mine drainage. In this mechanism, the initial carbon emission allowances are defined and allocated by local government for free or at auction or a combination of both [4,20] and coal mines decide their coal production under the emission allowance. Based on this mechanism, there are two kinds of decision makes in this problem. However, government only concerns the whole benefit and coal mines only consider their individual benefits. There is a conflict between government and coal mimes. Therefore, in order to solve this conflict, this paper adopts bi-level programming to balance the benefit between the whole and individuals. Meanwhile, this paper also considers the effect of mine water recycling on reducing waste water emission. Coal mines can have more coal production by recycling mine water to reduce waste water emission under the limited emission right.

Based on the above discussion, this paper proposes an optimization model that integrates a trade-and-cap mechanism and waste water recycling to mitigate mine water pollution. In Sect. 2, the government use to control pollution through emission allowance allocation, and the coal mines treat and reuse mine water to guarantee production under limited emission right. Then, as an abstraction of the real problem, a bi-level multi-objective model is built based on the discussion and a algorithm is applied in Sect. 3. In Sect. 4, a case study is presented to demonstrate the significance of the proposed model and solution method. Conclusions and future research directions are given in Sect. 5.

2 Problem Statement

In this paper, we consider a trade-and-cap mechanism to control the total waste water emission. Government decides the number of permits to discharge specific quantities of a specific mine water per time period. Coal mines need to decide their production plan according to permits allocated by government, because the quantity of mine water is related to the coal production. Meanwhile, coal mines can reduce their actual waste water emission though mine water recycling. In this problem, government hope to minimize the total waste water emission in the terms of ecology. But coal production is an important finance source for government. It's unreasonable to only consider ecology or economy singly. Therefore, government have to make a balance between economic and ecological benefits. And coal mines as profitable organizations only consider their economic benefits. The relationship between government and coal mines is shown as Fig. 1.

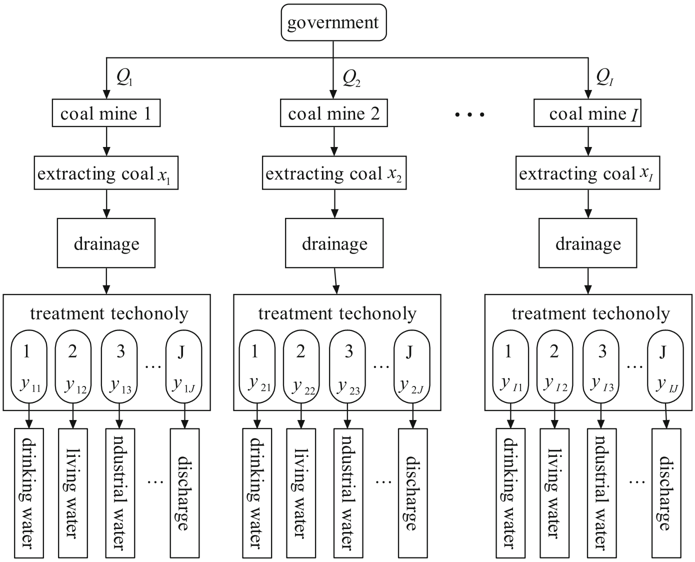


Fig. 1. The logic diagram of emission right allocation and mine water recycling

3 Modelling

Based on the analysis above, the hierarchical interests equilibrium between government and coal mines, and the economic and ecological conflicts are respectively the key points of solving the mine water emission reduction. Therefore, we adopt the bi-level programming to achieve the interests equilibrium which DMs in a hierarchical relationship to solve the emission right allocation, and the multi-objective programming to balance the economic and ecological benefits.

3.1 Notation

Index:

i: Coal mine index, $i = 1, 2, \cdots, I$;
j: Coal mine water treatment method index, $j = 1, 2, \cdots, J$, and the quality of treated water is from high to low when j from 1 to J, and method J can just make the mine water reach the emission standard.

Parameters:

ω: Tax rate of coal sale;
H: Maximum environment carrying capacity;
P^C: Unit coal price;
C_i^C: Unit coal mining cost of coal mine i;

B_{ij}^R: Unit economic benefit of reusing the waste water treated by method j in coal mine i;
C_{ij}^R: Unit economic cost of treating the waste water by method j in coal mine i;
θ_i: Coal drainage coefficient when coal mine i produces unit coal.

Decision variables:

Q_i: Waste water emission allowance of government allocating to coal mine i;
x_i: Coal production amount of coal mine i;
y_{ij}: Waste water amount of coal mine i treating by method j.

3.2 Objective for Government

(1) Economic Objective
For the region with abundant coal resource, coal industry is generally the main industry influencing the local economic development and human salary level. Therefore, local government must consider the economic objective, especially in developing country. Generally speaking, the financial revenue of the local government is determined by local tax revenue, central tax returns and transfer payments. The coal industry's contribution to regional financial revenue comes mainly from taxes. Therefore, the economic objective of government can be expressed as:

$$\max F_1 = \sum_{i=1}^{I} \left[(P^C - C^C) \times x_i \right] \times \omega, \tag{1}$$

(2) Ecological Objective
Environmental protection is a major responsibility for local government, and this is especially important in large scale coal mining regions. In terms of environmental protection, government hopes to minimize the actual waste water emission. Therefore, it is the optimal for ecological objective when total emission allowance is minimum.

$$\min F_2 = \sum_{i=1}^{I} Q_i. \tag{2}$$

3.3 Constraints for Government

(1) Maximum Environment Carrying capacity
Emission right allocation is to control the environmental damage effectively. Therefore, the total waste water should be under the maximum environment carrying capacity H. However, it is very difficult to estimate the exact value of the constraint due to the inherent uncertainty and complexity. Therefore, we adopt fuzzy number to express this kind of uncertainty.

$$\sum_{i=1}^{I} Q_i \leq H. \tag{3}$$

3.4 Objective for Coal Mine

For the market-based coal mines, pursuit of the highest profit is the absolute priority when the managers make decisions. Generally, coal mines' profit come mainly from coal production and sales. In this paper, we also consider the recycling and utilization of coal mine water. So the economic influence coming from waste water reuse and selling or buying emission right should be considered in the economic objective.

$$\max f_i = (P^C - C^C) \times x_i + \sum_{j=1}^{J} (B_{ij}^r - C_{ij}^r) \times y_{ij} - C_i^d \times (x_i \times \theta_i - \sum_{j}^{J} y_{ij}). \tag{4}$$

3.5 Constraints for Coal Mine

(1) Mining Ability Restrictions
Coal mines' mining ability is limited and can not be quickly improved. Therefore, when coal mine i makes a decision about the quantity that should mined, this quantity must not exceed its mining ability:

$$x_i \leq C_i^{\max}, \quad i = 1, 2, \cdots, I. \tag{5}$$

(2) Available Waste Water Restrictions
The amount of mine water drained by coal mine is relevant to the coal exacting amount, and the total amount of treated mine water cannot exceed the total mine water amount. Therefore, we can get

$$\sum_{j}^{J} y_{ij} = x_i \times \theta_i, \quad i = 1, 2, \cdots, I. \tag{6}$$

(3) Recycling Technology Restriction
To clean mine water to different quality, coal mine needs to use different methods. Obviously, the treated mine water amount by method j cannot exceed the maximum treatment capacity of this kine of technology the coal mine having. Therefore, the relationship can be described as:

$$0 \leq y_{ij} \leq T_{ij}, \quad i = 1, 2, \cdots, I, \quad j = 1, 2, \cdots, J. \tag{7}$$

(4) Recycling Water Demand Restriction
The treated mine water will be used in different method according to their qualities respectively, but users' demands for water are limited. Therefore, treated mine water cannot exceed the user's demand, otherwise, the excessive water will be wasted.

$$0 \leq y_{ij} \leq D_{ij}, \quad i = 2, \cdots, I, \quad j = 1, 2, \cdots, J - 1. \tag{8}$$

(5) Environmental Protection Restriction
To protect water environment, the actual waste water emission cannot exceed the allowance allocated by government. And the actual waste water emission should be the waste water without recycling. So the relationship is as following:

$$x_i \times \theta_i - \sum_{j}^{J} y_{ij} \le Q_i, \quad i = 1, 2, \cdots, I. \tag{9}$$

3.6 Global Model

Based on the previous discussion, by integrating Eqs. (1)–(9), the following global expected model of bilevel programming can now be formulated for the problem:

$$\begin{aligned}
&\max F_1 = \sum_{i=1}^{I} \left[(P^C - C^C) \times x_i \right] \times \omega \\
&\min F_2 = \sum_{i=1}^{I} Q_i \\
&s.t. \begin{cases}
\sum_{i=1}^{I} Q_i \le H \\
Q_i \ge 0, \quad i = 1, 2, \cdots, I \\
\max f_i = (P^C - C^C) \times x_i + \sum_{j=1}^{J} (B_{ij}^r - C_{ij}^r) \times y_{ij} \\
\qquad\qquad - C_i^d \times (x_i \times \theta_i - \sum_j^J y_{ij}) \\
\begin{cases}
x_i \le C_i^{\max}, \quad i = 1, 2, \cdots, I \\
x_i \times \theta_i \le \sum_{j=1}^{J} T_{ij}^r + T^d \quad i = 1, 2, \cdots, I \\
\sum_j^J y_{ij} \le x_i \times \theta_i, \quad i = 1, 2, \cdots, I \\
0 \le y_{ij} \le T_{ij}, \quad i = 1, 2, \cdots, I, \quad j = 1, 2, \cdots, J \\
0 \le y_{ij} \le D_{ij}, \quad i = 2, \cdots, I, \quad j = 1, 2, \cdots, J-1 \\
x_i \times \theta_i - \sum_j^J y_{ij} \le Q_i, \quad i = 1, 2, \cdots, I.
\end{cases}
\end{cases}
\end{aligned} \tag{10}$$

4 Solution Approach

In this section, we adopt a KKT to solve the bi-level problem and fuzzy goals programming to solve multi-objective problem.

4.1 Fuzzy Goal Programming for Bi-Level Multi-objective Model

The basic concept of fuzzy goal programming approaches is that each lower level DM considers the upper level DM's goal when seeking to optimize their own objective function. In the decision process, the fuzzy goal membership functions for all the DMs' variables are considered and the fuzzy programming problem is solved using a constraint on the overall satisfactory degree of any upper levels. [13,14]. As membership levels are in the same dimension, this method can also solve the weighted sum scalarization with different dimensions in model (10)

Meanwhile, this method has been previously adopted to solve multi-objective model problems in different areas [1,10] and has proved to be highly effective. Therefore, this paper adopts fuzzy goal programming to transform the bi-level multi-objective model into a single level and single objective model.

First, the optimal solution for each objective of each decision maker f_ξ^B is calculated and the DMs decided on the worst value f_ξ^W they can accept (ξ is the objective index), from which the fuzzy goal $f_1(x,y,z) \gtrsim f_1^B$ and $f_2(x,y,z) \gtrsim f_2^B$ can be determined, and their membership functions are as Eq. (11)

$$\mu_\xi(f_\xi(q,x,y)) = \begin{cases} 1, & \text{if } f_\xi(q,x,y) \le f_\xi^B, \\ \frac{f_\xi(q,x,y)-f_\xi^W}{f_\xi^B-f_\xi^W}, & \text{if } f_\xi^B < f_\xi(q,x,y) \le f_\xi^W, \\ 0, & \text{if } f_\xi(q,x,y) \ge f_\xi^W. \end{cases} \tag{11}$$

Using Eq. (11), the multi-objective model can be transformed into a single objective model by weighting and summing the negative deviations between the actual satisfactory level and the aspiration satisfactory level for each objective, as shown in model (12):

$$\begin{aligned} &\max F = \sum_{\xi\in\Xi} \omega_\xi \times \mu_\xi(F_\xi(q,x,y)) \\ &s.t. \begin{cases} \bar{\mu}_{ij} - \mu_{ij}(f_{ij}(q,x,y)) \le \hat{\sigma} \\ (q,x,y) \in S, \end{cases} \end{aligned} \tag{12}$$

where the auxiliary variable ω_ξ is the importance degree for objective ξ and $\sum_{\xi\in\Xi}\omega_\xi = 1$; $\bar{\mu}_{ij}$ is the individual aspiration levels of each coal mine; $\mu_{ij}(f_{ij}(q,x,y)$ is the actual the practical satisfaction level of each decision maker; $\hat{\sigma}$ is the maximum deviation between the practical satisfaction level and the aspiration satisfaction level of each lower decision maker; S is the feasible region of Eq. (10). By minimizing the weighted sum of these deviation, decision makers can determine an overall satisfactory solution close to the optimal solution.

However, based on these characteristics, however, there is a limitation to the use of interactive fuzzy programming as it can only be used to solve decision making problems in organizations with cooperative behavior. And the solution of model (12) is not optimal for lower decision makers. To solve this problem, we adopt KKT to transform the lower model to guarantee the optimal solution of lower decision makers.

4.2 KKT

Karush-Kuhn-Tucker (KKT) approach has been proven to be a valuable analysis tool with a wide range of successful applications for bi-level programming [11]. The fundamental strategy for the KKT approach is that it replaces the follower's problem with its KKT conditions and appends the resultant system to the leader's problem [8,12]. Based on previous work by Shi et al. [15], we transformed the bi-level model into a single-level model. However, KKT cannot guarantee the equality between lower decision makers when there are multiple lower decision makers in a bi-level problem. But in this paper, we have used fuzzy goal programming to transform the bi-level model into single model and can guarantee the benefit balance between lower decision makers by the satisfactory deviation. Therefore, in this paper, we just adopt KKT to seek the optimal solution of lower model under upper decision maker's strategy. And the model (10) is finally transformed as Eq. (13).

$$
\begin{aligned}
&\max F = \omega_1 \times \mu_\xi(F_1(q,x,y)) + \omega_2 \times \mu_\xi(F_2(q,x,y)) \\
&s.t. \begin{cases}
F_1 = \sum_{i=1}^{I} \left[(P^C - C^C) \times x_i\right] \times \omega \\
F_2 = \sum_{i=1}^{I} Q_i \\
\mu_\xi(f_\xi(q,x,y)) = \begin{cases} 1, & \text{if } f_\xi(q,x,y) \le f_\xi^B, \\ \frac{f_\xi(q,x,y) - f_\xi^W}{f_\xi^B - f_\xi^W}, & \text{if } f_\xi^B < f_\xi(q,x,y) \le f_\xi^W, \\ 0, & \text{if } f_\xi(q,x,y) \ge f_\xi^W, \end{cases} \\
\bar{\mu}_{ij} - \mu_{ij}(f_{ij}(q,x,y)) \le \hat{\sigma} \\
\sum_{i=1}^{I} Q_i \le H \\
Q_i \ge 0, \quad i = 1,2,\cdots,I \\
x_i \le C_i^{\max}, \quad i = 1,2,\cdots,I \\
x_i \times \theta_i \le \sum_{j=1}^{J} T_{ij}^r + T^d \quad i = 1,2,\cdots,I \\
\sum_j^J y_{ij} \le x_i \times \theta_i, \quad i = 1,2,\cdots,I \\
0 \le y_{ij} \le T_{ij}, \quad i = 1,2,\cdots,I, \quad j = 1,2,\cdots,J \\
0 \le y_{ij} \le D_{ij}, \quad i = 2,\cdots,I, \quad j = 1,2,\cdots,J \\
x_i \times \theta_i - \sum_j^J y_{ij} \le Q_i, \quad i = 1,2,\cdots,I \\
(P^C - C^C) - C_i^d \times \theta_i + v_1 + (v_2 - v_3 + v_{11}) \times \theta_i = 0 \\
\sum_{j=1}^{J} (B_{ij} - C_{ij}) + C_i^d + v_3 + v_4 - v_5 + v_6 - v_{11} = 0 \\
v_1 g_1(Q_i,x_i,y_{ij}) + v_2 g_2(Q_i,x_i,y_{ij}) + v_3 g_3(Q_i,x_i,y_{ij}) \\
+v_4 g_4(Q_i,x_i,y_{ij}) + v_5 g_5(Q_i,x_i,y_{ij}) + v_6 g_6(Q_i,x_i,y_{ij}) \\
+v_7 g_7(Q_i,x_i,y_{ij}) = 0 \\
g_1(Q_i,x_i,y_{ij}) = C_i^{\max} - x_i \ge 0, \quad i = 1,2,\cdots,I \\
g_2(Q_i,x_i,y_{ij}) = \sum_{j=1}^{J} T_{ij}^r + T^d - x_i \times \theta_i \ge 0, \quad i = 1,2,\cdots,I \\
g_3(Q_i,x_i,y_{ij}) = x_i \times \theta_i - \sum_j^J y_{ij} \ge 0, \\
g_4(Q_i,x_i,y_{ij}) = T_{ij} - y_{ij}, \quad i = 1,2,\cdots,I, \quad j = 1,2,\cdots,J \\
g_5(Q_i,x_i,y_{ij}) = y_{ij} \ge 0, \quad i = 1,2,\cdots,I, \quad j = 1,2,\cdots,J \\
g_6(Q_i,x_i,y_{ij}) = D_{ij} - y_{ij} \ge 0, \quad i = 2,\cdots,I, \quad j = 1,2,\cdots,J \\
g_7(Q_i,x_i,y_{ij}) = Q_i - x_i \times \theta_i + \sum_j^J y_{ij} \ge 0.
\end{cases}
\end{aligned}
\tag{13}
$$

5 Case Study

5.1 Case Presentation

Lu'an coal field is located on the southeastern edge of the middle eastern Qinshui coal field, which is in the center of Shanxi Province. There are nine major collieries (i.e., $i = 4$) in this area. These are the Shiyijie mine, the Changcun mine, the Zhangcun mine and the Wangzhuang mine.

In Lu'an coal field, there is average annual rainfall of 584 mm, and an average annual evaporation of 1732 mm. With the capacity expansion of the coal mines, coal mining quantity has significantly increased, and this caused mine drainage increases simultaneously. Moreover, untreated mine water damages the surrounding water environment. Meanwhile, due to the water from the mine drainage, the groundwater supplement is insufficient, which affects plant growth and adds to environmental degradation. Therefore, this paper adopt Lu'an coal field as a case to show the practicality of the model.

5.2 Data Collection

Detailed data for the research region were obtained from the Statistical Yearbook of Chinese coal industry (National Bureau of Statistics of the Peoples Republic of China (2013)), the Statistical Yearbook of Chinese energy (National Bureau of Statistics of the Peoples Republic of China (2013)) and field research. Certain parameters taken from the Statistical Yearbooks and data published by the companies are shown in Tables 1, 2 and 3.

5.3 Results and Discussions

First, the upper and lower bounds of the government and coal mines objective functions were calculated to identify the membership function for the fuzzy goals, as shown in Table 4.

Then, we can build the fuzzy goals though the optimal solutions and worst solutions of each objective of government and coal mines. And we set $\omega_1 = \omega_2 = 0.5$, $\bar{\mu}_1 = \bar{\mu}_2 = \bar{\mu}_3 = \bar{\mu}_4 = 1$. Finally, through adapting the maximum deviation $\hat{\sigma}$, we can get some groups optimal solutions under different attitude of government, as Table 5.

Table 1. Parameters for government

Government	H (10^4 t)	ω
	1000	0.17

Table 2. Parameters of coal mine (1)

	B_{ij}^r (RMB/m^3)			C_{ij}^r (RMB/m^3)			T_{ij}^r (10^4 m^3)			D_{ij}^r (10^4 m^3)		
	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$
$i=1$	2.8	2.8	4.1	2	1.7	1.5	5	20	120	26	38.5	359
$i=2$	2.8	2.8	4.1	0	1.66	1.49	0	15	90	20	28.1	240
$i=3$	2.8	2.8	4.1	0	1.72	1.52	0	10	75	17	30	305
$i=4$	2.8	2.8	4.1	1.98	1.59	1.4	3	34	100	24.9	35.7	367

Table 3. Parameters of coal mine (2)

	$C_i^{\max}$ (10^4 t)	P_i^c (yuan/t)	C_i^c (yuan/t)	θ_i (m^3/t)	T_i^d (10^4 m^3)	C_i^d (RMB/m^3)
$i = 1$	230	780	219	2.15	500	0.63
$i = 2$	170	780	230	2.12	320	0.6
$i = 3$	136	780	227	2.14	283	0.57
$i = 4$	250	780	220	2.12	490	0.61

Table 4. Pay-off table with 5 individual single-objective problems

	max F_1	min F_2	max f_1	max f_2	max f_3	max f_4
F_1	68914.81	29885.33	44858.20	30641.72	28683.95	45683.80
F_2	1000	0	1000	1000	1000	1000
f_1	128173.37	60852.47	169949.24	12156.43	12156.43	12156.43
f_2	94745.76	27240.28	62580.98	77884.18	62580.98	62580.98
f_3	75799.63	32076.54	55084.51	55084.51	77157.82	55084.51
f_4	140199.70	57563.49	18477.52	18477.52	18477.52	132121.31

Table 5. Optimal solutions under different attitude of government

		Q_i (10^4 m^3)	x_i (10^4 t)	y_{ij} (10^4 m^3)			f_i (10^4 RMB)	$\mu(f_i)_i$	F_1 (10^4 RMB)	F_2 (10^4 m^3)	$\mu(F)$
				$j = 1$	$j = 2$	$j = 3$					
$\hat{\sigma} = 1$	$i = 1$	0.00	107.84	5.00	20.00	250.00	60852.47	0.00	61846.81	810.00	0.5044
	$i = 2$	320.00	182.77	0.00	15.00	100.00	102073.49	1.00			
	$i = 3$	0.00	56.82	0.00	10.00	140.00	32076.55	0.00			
	$i = 4$	490.00	304.55	3.00	34.00	210.00	170248.06	1.00			
$\hat{\sigma} = 0.8$	$i = 1$	100.00	147.06	5.00	20.00	250.00	82671.82	0.20	67600.55	966.60	0.4998
	$i = 2$	320.00	182.77	0.00	15.00	100.00	102073.49	1.00			
	$i = 3$	56.60	78.26	0.00	10.00	140.00	44007.47	0.20			
	$i = 4$	490.00	304.55	3.00	34.00	210.00	170248.06	1.00			
$\hat{\sigma} = 0.6$	$i = 1$	200.00	186.27	5.00	20.00	250.00	104491.18	0.40	68525.06	1000.00	0.495
	$i = 2$	320.00	182.77	0.00	15.00	100.00	102073.49	1.00			
	$i = 3$	113.20	99.70	0.00	10.00	140.00	55938.39	0.40			
	$i = 4$	366.80	253.64	3.00	34.00	210.00	141915.94	0.75			
$\hat{\sigma} = 0.4$	$i = 1$	300.00	225.49	5.00	20.00	250.00	126310.53	0.60	68085.13	1000.00	0.4894
	$i = 2$	236.20	147.56	0.00	15.00	100.00	82476.55	0.74			
	$i = 3$	169.80	121.14	0.00	10.00	140.00	67869.30	0.60			
	$i = 4$	294.00	223.55	3.00	34.00	210.00	125174.24	0.60			
$\hat{\sigma} = 0.38$	$i = 1$	310.00	229.41	5.00	20.00	250.00	128492.46	0.62	68029.88	1000.00	0.4887
	$i = 2$	210.74	136.87	0.00	15.00	100.00	76522.63	0.66			
	$i = 3$	175.46	123.28	0.00	10.00	140.00	69062.40	0.62			
	$i = 4$	303.80	227.60	3.00	34.00	210.00	127427.93	0.62			

From Table 5, when the maximum deviation $\hat{\sigma} = 1$, which means government doesn't consider the equality between coal mines fully, government can get its optimal solution and the satisfactory is 0.5044. When government begin to decrease the maximum deviation from 1 to 0.38 to consider the equality between coal mines, its satisfactory level drop continuously. Because coal mines only concern about economic benefit, the ecological benefit decreases faster than economic benefit with $\hat{\sigma}$ draping. $\hat{\sigma}$ from 1 to 0.6, the ecological objective value drop to the worst value. And $\hat{\sigma}$ from 0.6 to 0.38, the whole economic benefit decreases and the benefit balance between coal mines is improved. When $\hat{\sigma} < 0.38$, there is no feasible solution for model (13).

Comparing these results, we can find government prefer allocate emission right to these coal mines with better waste water recycling capacity because which can guarantee the economic and ecological benefit simultaneously in the biggest extend. And when government consider the whole satisfactory level, the upper benefit would be damaged. Therefore, government should push these coal mines faster the reformation step to achieve the economy and ecology development together.

6 Conclusion

This paper proposed a bi-level multi-objective programming model to deal with economic and ecological conflicts in large scale coal fields. An environmental protection based mining quotas competition mechanism was established using the proposed model which considers not only the relationship between the government and coal mines, but also the equilibrium between economic development and environmental protection. To solve the complex bi-level multi-objective programming model and an extended KKT and fuzzy goal programming approach were combined as a solution method. The proposed method was then applied to the Lu'an coal field which includes four major coal mines. By inputting the data into the model and computing it using the proposed solution approach, the effectiveness of the model was demonstrated.

References

1. Baky IA (2010) Solving multi-level multi-objective linear programming problems through fuzzy goal programming approach. Appl Math Model 34(9):2377–2387
2. Clo S (2009) The effectiveness of the EU emissions trading scheme. Clim Policy 9(3):227–241
3. Coase RH (1960) The problem of social cost. J Law Econ 3:1–44
4. Cong RG, Wei YM (2012) Experimental comparison of impact of auction format on carbon allowance market. Renew Sustain Energ Rev 16(6):4148–4156
5. Gersbach H, Requate T (2004) Emission taxes and optimal refunding schemes. J Public Econ 88(3):713–725
6. Goulder LH, Hafstead MA, Dworsky M (2010) Impacts of alternative emissions allowance allocation methods under a federal cap-and-trade program. J Environ Econ Manag 60(3):161–181
7. Haibin L, Zhenling L (2010) Recycling utilization patterns of coal mining waste in China. Resour Conserv Recycl 54(12):1331–1340
8. Hanson MA (1981) On sufficiency of the Kuhn-Tucker conditions. J Math Anal Appl 80(2):545–550

9. Harrington W, Morgenstern RD (2007) Economic incentives versus command and control: whats the best approach for solving environmental problems? In: Acid in the Environment. Springer, pp 233–240
10. Hu CF, Teng CJ, Li SY (2007) A fuzzy goal programming approach to multi-objective optimization problem with priorities. Eur J Oper Res 176(3):1319–1333
11. Lu J, Shi C, Zhang G (2006) On bilevel multi-follower decision making: general framework and solutions. Inf Sci 176(11):1607–1627
12. Lu J, Shi C, Zhang G, Dillon T (2007) Model and extended Kuhn-Tucker approach for bilevel multi-follower decision making in a referential-uncooperative situation. J Glob Optim 38(4):597–608
13. Moitra Bhola Nath, Pal Bijay Baran (2002) A fuzzy goal programming approach for solving bilevel programming problems. In: Advances in Soft Computing, AFSS 2002. Springer, Heidelberg, pp 91–98
14. Pal BB, Moitra BN (2003) A fuzzy goal programming procedure for solving quadratic bilevel programming problems. Int J Intell Syst 18(5):529–540
15. Shi C, Lu J, Zhang G (2005) An extended Kuhn-Tucker approach for linear bilevel programming. Appl Math Comput 162(1):51–63
16. Silva LF, de Vallejuelo SFO, Martinez-Arkarazo I et al (2013) Study of environmental pollution and mineralogical characterization of sediment rivers from Brazilian coal mining acid drainage. Sci Total Environ 447:169–178
17. Sun W, Wu Q, Dong D, Jiao J (2012) Avoiding coal-water conflicts during the development of Chinas large coal-producing regions. Mine Water Environ 31(1):74–78
18. Wolkersdorfer C, Bowell R (2005) Contemporary reviews of mine water studies in Europe, Part 2. Mine Water Environ 24(1):2–37
19. Younger PL, Wolkersdorfer C (2004) Mining impacts on the fresh water environment: technical and managerial guidelines for catchment scale management. Mine Water Environ 23:s2–s80
20. Zhixin Z, Ya L (2011) The impact of carbon tax on economic growth in China. Energ Procedia 5:1757–1761

Modeling and Solving the Vehicle Routing Problem with Multiple Fuzzy Time Windows

Fang Yan(✉) and Yuanyuan Wang

School of Economic and Management, Chongqing Jiaotong University,
Chongqing 400074, People's Republic of China
yanfang@cqjtu.edu.cn

Abstract. Considering that the customers could accept service in multiple time periods and the fuzziness of service time periods, this paper deals with the multiple time windows as fuzzy variables and quantifies the customer satisfaction level according to the membership function of the beginning time to be served, on the basis of the given acceptable satisfaction level, the vehicle routing model with multiple fuzzy time windows is constructed in order to minimize the transportation cost and number of vehicles and maximize the satisfaction level. Then according to the model characteristics, we use the punishment factors to deal with the constraints and apply particle swarm operations to solve the proposed problems. The experimental results show the effectiveness of proposed algorithm in solving the vehicle routing problems with multiple fuzzy time windows. Comparing the calculated results with the hard time window model results, it is found that our proposed model is more effective to reduce the cost of distribution.

Keywords: Multiple fuzzy time windows · Particle swarm operations · Vehicle routing · Customer satisfaction

1 Introduction

Vehicle routing problem was first proposed by Dantzing and Ramser [5]. After more than 50 years of research, it has attracted wide attention of scholars in the field of operations research and combinatorial optimization. The vehicle routing problem with time window (VRPTW) is add the time window constraint on the basis of the capacity constraint model. At present, domestic and foreign scholars have done a lot of research on the VRPTW problem, Desrochers et al. [6] solved VRPTW problem by combined branch-and-bound algorithm and column generation algorithm. Bent and Hentenryck [3] proposed a two-stage hybrid algorithm for VRPTW problem. In addition, the algorithm for the VRPTW problem is also includes genetic algorithm [15], particle swarm optimization algorithm [1], ant colony system [17], etc. However, considering the customer service time requirements are not completely rigid, Lin [12] considered customer satisfaction

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_69

level and proposed vehicle routing problems based on the fuzzy time window. Ghannadpour et al. [9] presented a multi-objective dynamic vehicle routing problem with fuzzy time windows considered customer-specific time windows and proposed the genetic algorithm (GA) and three basic modules to solve the problem. While these studies have contributed significantly to solving the vehicle routing problem with single time window, customers usually have more than one time period to receive service, they ignored the vehicle routing problem with multiple time windows (VRPMTW).

The vehicle routing problem with multiple time windows (VRPMTW) is an extension problem of the traditional vehicle routing problem, refers to each customer has a number of non overlapping time windows to receive services, delivery vehicles must choose one time window for service, and compared with the single time window of vehicle routing problem, VRPMTW is much closer to the reality. The research on VRPMTW is mainly focused on two aspects: model and algorithm. In the aspect of model, Doerner et al. [7] presented the model of vehicle routing problem with multiple interdependent time windows and solved the problem by several variants of a heuristic constructive procedure as well as a branch-and-bound based algorithm. Ma et al. [14] considered delivery can be splited and established the split delivery vehicle routing problem with multiple time windows (SDVRPMTW) model. Yan et al. [18] applied VRPMTW model in military oil transport path optimization and solved problem with particle swarm optimization algorithm. And the research of algorithms mainly focus on the ant colony system algorithm [8], simulated annealing algorithm [13], hybrid intelligent algorithm [4], hybrid variable neighborhood-tabu search heuristic [2] and intelligent water drop algorithm [11], etc. However, customer usually have multiple fuzzy time windows and service time requirements are not completely rigid in reality. This kind of problem is more complicated and more realistic than VRPMTW, so it is necessary to carry out relevant research.

In this paper, VRPMTW in a fuzzy environment is considered and customer satisfaction is quantified by the time to start service. We aim to minimize transportation cost, minimize vehicles, and maximum customer satisfaction, thus establish the model of vehicle routing problem with multiple fuzzy time windows. The particle swarm optimization algorithm is used to solve the problem and the experimental results are analyzed and discussed.

2 Problem Statement and Modeling

2.1 Problem Statement

The vehicle routing problem with multiple fuzzy time windows can be stated as: A distribution center has m cars to serve n customers, customer i has W_i non overlapping fuzzy time windows. Each customer's needs, the maximum loading capacity of each vehicle and the distance between any two customers is known. Vehicles starting from the distribution center, select one time window of customer to serve and return distribution center until completion of distribution.

On the basis of satisfying customer demand, vehicle loading capacity and customer time window, the objective function is optimized through the reasonable path planning. The model considers the following constraints: (1) Vehicle loading capacity constraint: the actual loading of each vehicle must not exceed the maximum loading capacity. (2) Multi fuzzy time windows constraint: each customer has multiple fuzzy time windows, but vehicle can only choose a time window service. (3) Customer satisfaction constraint: customer satisfaction is greater than the value of the decision maker. (4) Visit uniqueness constraint: each customer can only served by a car and can only serve once. (5) Central constraint: vehicle start from the distribution center and after implement the customer service to return. (6) Loop elimination constrain: the path of each vehicle can only exist the loop between the starting point and point, there can be no other circuit.

2.2 Notations

The variables and parameters used in the model are:

$L = \{1, 2, \cdots, n\}$	: customer set, with 0 and $n+1$ said distribution center
$K = \{1, 2, \cdots, m\}$	: vehicle set;
Q_k	: the loading capacity of vehicle k;
C	: vehicle start-up costs;
q_i	: the demand of customer i;
D_k	: the longest running distance of vehicle k;
c_{ij}	: the cost of traveling from nodes i to j;
t_i	: the service start time of customer i ;
s_i	: the service duration time of customer i ;
t_{ij}	: the travel time from customer to customer i to j ;
W_i	: the number of customer time windows;
$[E_i^\alpha, a_i^\alpha, b_i^\alpha, L_i^\alpha]$	: the fuzzy time window α of customer i;
E_i^α	: the endurable earliness time of time window α of customer i;
a_i^α	: the earliest start service time which customer i expect to be served in time window α;
b_i^α	: the latest start service time which customer i expect to be served in time. window α;
x_{ijk}	: a binary variable, $x_{ijk} = 1$, if vehicle travels from customer i to customer j, otherwise, $x_{ijk} = 0$;
y_i^α	: a binary variable, $y_i^\alpha = 1$, if vehicle serve the time window α of customer i, otherwise, $y_i^\alpha = 0$.

2.3 Fuzzy Time Window

Trapezoidal fuzzy time windows is used in this paper, as shown in Fig. 1. Each customer can own multiple trapezoidal fuzzy time windows, customer is satisfied when be served in the expected time period and the satisfaction level is 1, while customer satisfaction will be reduced when the service time is not met or exceeded the customer expected service time.

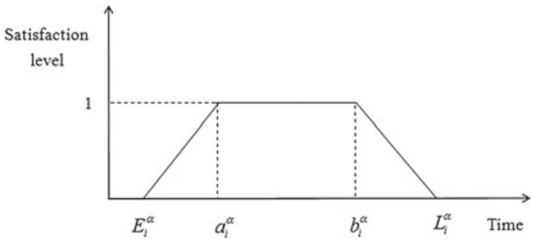


Fig. 1. Trapezoidal fuzzy time window

When the customer has multiple trapezoidal fuzzy time window can accept the service, customer satisfaction can be defined by the membership function of the service start time:

$$\mu_i(t_i) = \begin{cases} 0, t_i < E_i^\alpha \\ (t_i - E_i^\alpha)/(a_i^\alpha - E_i^\alpha), E_i^\alpha < t_i < a_i^\alpha \\ 1, a_i < t_i < b_i \\ (L_i^\alpha - t_i)/(L_i^\alpha - b_i^\alpha), b_i^\alpha < t_i < L_i^\alpha \\ 0, t_i > L_i^\alpha. \end{cases} \tag{1}$$

2.4 Modelling

The vehicle routing problem with multiple fuzzy time windows is as follows:

$$\min \ Z_1 = \sum_{i=0}^{n} \sum_{j=0}^{n+1} \sum_{k=1}^{m} c_{ij} x_{ijk} \tag{2}$$

$$\min \ Z_2 = C \sum_{j=0}^{n} x_{0jk}, \tag{3}$$

$$\max \ Z_3 = \frac{1}{n} \sum_{i=1}^{n} \mu_i(t_i), \tag{4}$$

$$\text{s. t. } \mu_i(t_i) \geq \eta_i, \forall i \in L, \tag{5}$$

$$\sum_{i=1}^{n} \left(q_i \sum_{j=0}^{n} x_{ijk} \right) \leq Q_k, \forall k \in K, \tag{6}$$

$$\sum_{i=0}^{n} \sum_{j=1}^{n+1} d_{ij} x_{ijk} \leq D_k, \forall k \in K, \tag{7}$$

$$\sum_{j=1}^{n+1} x_{0jk} = 1, \forall k \in K, \tag{8}$$

$$\sum_{i=1}^{n+1} x_{in+1k} = 1, \forall k \in K, \tag{9}$$

$$\sum_{i=1}^{n}\sum_{k=1}^{m} x_{ijk} = \sum_{i=1}^{n}\sum_{k=1}^{m} x_{jik} = 1, \forall j \in L, \tag{10}$$

$$\sum_{i,j \in S \times S} x_{ijk} \leq |S| - 1, S \subseteq L, \forall k \in K, \tag{11}$$

$$L_i^{\alpha} \leq E_i^{\alpha+1}, \forall i \in L, \alpha \in \{1, 2, \cdots W_i - 1\}, \tag{12}$$

$$\max \left\{ \sum_{\alpha=1}^{W_i} y_i^{\alpha} E_i^{\alpha}, (t_i + s_i + t_{ij})\, x_{ijk} \right\} \leq t_j, \forall i, j \in L, \forall k \in K, \tag{13}$$

$$t_j \leq \sum_{\alpha=1}^{W_i} y_j^{\alpha} L_j^{\alpha}, \forall j \in L, \tag{14}$$

$$\sum_{\alpha=1}^{W_i} y_i^{\alpha} = 1, \forall i \in L, \tag{15}$$

$$x_{ijk} = 0 \text{ or } 1 \forall i, j, k, \tag{16}$$

$$y_i^{\alpha} = 0 \text{ or } 1, \forall i \in L, \alpha \in \{1, 2, \cdots, W_i\}. \tag{17}$$

Objective (2) minimizes the total travel cost of all vehicles which is the most important objective to the decision makers. Objective (3) minimums the number of vehicles. Objective (4) maximizes the average customer satisfaction. Equation (5) ensures that each customer satisfaction is higher than η_i, η_i is given by the decision maker based on experience. Equation (6) ensures that the loading of each vehicle is not more than the maximum loading capacity. Equation (7) makes sure the distance of each vehicle is not more than the longest running distance. Equations (8) to (10) ensure that each customer is served by a car and the vehicle flow keep balance. Equation (11) is to eliminate the sub loop. Equation (12) is to sort time windows by time. Equations (13) and (14) makes sure that customer is served in time window. Equation (15) ensures each customer has a time window to be served. Equations (16) and (17) is the variable range.

3 Particle Swarm Optimization

Particle swarm optimization algorithm was first proposed by Kennedy and Eberhart [10], and it is a kind of stochastic optimization technique based on population. Compared with other evolution algorithm based on population, their initialization is a set of random solutions and through the iterative to search the optimal solution. The difference is that evolutionary computation follows the principle of survival of the fittest and PSO simulates the society. PSO express each possible solution as a particle of group, each particle has its own position vector and velocity vector, as well as a fitness which is determined by the objective function. All particles are flying at a certain speed in the search space, by following the optimal value of the current search to find the global optimal value.

3.1 Structure of Solution

Construct the expression of the solution and make the solution correspond to the particles in the PSO algorithm is the key link of the algorithm. Salman et al. [16] described the expression of solution in detail, we draw lessons from this method. A $3L$-dimensional matrix is used to represent the solution of the proposed model with L customers, each customer corresponds to three dimensional: vehicle number vector, vehicle ranking vector and time windows number vector. In order to express clearly and calculate conveniently, the $3L$-dimensional matrix of each particle is divided into three L-dimensional vectors, they are X_v (vehicle number vector), X_r (vehicle ranking vector) and X_w (time windows number vector). Corresponding velocity vectors V are X_v, X_r and X_w.

For example, there are three vehicles to complete the distribution tasks of eight customers, the position vector of a particle is: Due to the original assumption that the vehicle starting from the distribution center and finally return to

Consumers	1	2	3	4	5	6	7	8
X_v	1	1	1	2	3	3	3	3
X_r	1	3	2	1	4	1	3	2
X_w	1	2	1	1	1	1	2	1

the distribution center, the corresponding vehicle routing is:

Vehicle 1	$0 \to 1 \to 3 \to 2 \to 0$
Vehicle 1	$0 \to 4 \to 0$
Vehicle 1	$0 \to 6 \to 8 \to 7 \to 5 \to 0$

In the coding, the 0 omitted in order to simplify the calculation. The third line represents the time window number of each customer to be served, customer 1 choose the first time window service, for the customer 2 choose second time window service and for customers 3 choose the first time window service, etc.

3.2 The Fitness Value Function

According to the objective function (2), (3), (4) and the constraint conditions design the fit value function, as shown in the following

$$\min Z = Cm + \sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=1}^{m} c_{ij}x_{ijk} + \beta_1 \sum_{k=0}^{m} \max\left\{\left[\sum_{i=1}^{n}\left(d_i \sum_{j=0}^{n} x_{ijk}\right) - Q_k\right], 0\right\}$$
$$+ \beta_2 \sum_{i=1}^{n} \max(E_i^{\alpha} - t_i, 0) + \beta_3 \sum_{i=1}^{n} \max(t_i - L_i^{\alpha}, 0) \tag{18}$$
$$+ \beta_4 \sum_{i=1}^{n} \max\left[\eta_i - \mu_i\left(t_i\right), 0\right],$$

where β_1, β_2, β_3 and are punishment coefficients, β_1 is the penalty factor for overloading, β_2 and β_3 penalty coefficient for violation of the time windows, β_4 is the penalty coefficient for violation of the customer satisfaction. Decision makers can set the punishment coefficient base on their predilection, thus eliminating the infeasible solution.

3.3 Particle Updating

Suppose there are n particles in the D-dimensional search space, the position vector of the particle i is $X_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$, the optimal position of the particle i during the flight is, global best position is $P_{i\text{best}} = (p_{i1}, p_{i2}, \cdots, p_{iD})$, $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$ is the flight speed of particle. Using the update method of Li et al. [11] to update, the update formula is as follows:

$$v_{id}\left(k+1\right) = \omega \times v_{id}\left(k\right) + c_1 \times \text{rand}\left(\right) \times \left[p_{id}\left(k\right) - x_{id}\left(k\right)\right] + c_2 \times \text{rand}\left(\right) \times \left[p_{gd}\left(k\right) - x_{id}\left(k\right)\right], \tag{19}$$

$$x_{id}(k+1) = x_{id}\left(k\right) + v_{id}\left(k+1\right) 1 \quad \leq i \leq n; 1 \leq d \leq D. \tag{20}$$

w is the inertia factor which is calculated by the following formula:

$$\omega(\text{t}) = \omega(\text{gen}) + (\text{t} - \text{gen})/(1 - \text{gen}) \times [\omega(1) - \omega(\text{gen})], \tag{21}$$

where c_1 and c_2 are accelerating factor, rand () generate random number between 0 and 1.

3.4 The Procedure of PSO

Step 1. Randomly generated the initial position and velocity of the particles within the range of the values.
Step 2. Based on formula (18) to calculate the fitness value of each particle.
Step 3. Calculate the fitness value of each particle and compared with the fitness value of the best position it has experienced, if the fitness value of the current location is better, set the current position as its personal best.
Step 4. Compare the personal best fitness of each particle with global best fitness, if the fitness value of the current location is better, set the current position as its global best.
Step 5. Update the particle velocities and positions according to the formulas (19) and (20).
Step 6. If the stopping criterion is met, stop; otherwise, return to Step 2.

4 A Case Study

In this section, we use an example to illustrate the model of vehicle routing problem with multiple fuzzy time windows. Example is as follows: a distribution center with six vehicles to provide services for 15 customers and each customer has two fuzzy time windows, the maximum loading capacity of each vehicle is 50 tons and the furthest distance is 120 km, the average vehicle speed is 35 km/h with the driving cost is 5 yuan/km and the start-up cost of each vehicle is 100 yuan, the average customer satisfaction level is set to 0.7 (Table 1).

Table 1. Customer information

Customer	d_i	s_i	Coordinate	E_i^1	a_i^1	b_i^1	L_i^1	E_i^2	a_i^2	b_i^2	L_i^2
1	4	0.3	(1,8)	6.5	8	9	10.5	12.5	14	15.6	17.1
2	9	0.4	(9,−8)	7	8.5	9.5	11	11.5	13	15	16.5
3	6	0.4	(−16,−20)	7	8.5	9.8	11.3	12	13.5	14.6	16.1
4	2	0.2	(19,12)	7	8.5	9.5	11	11.5	13	14.8	16.3
5	9	0.3	(−21,−15)	7	8.5	10	11.5	12	13.5	14	15.5
6	3	0.5	(3,25)	7	8.5	10.4	11.4	11.5	13	14	15.5
7	5	0.3	(12,1)	6.5	8	9	10.5	11.5	13	14	15.5
8	7	0.2	(14,−19)	7	8.5	9.5	11	11.3	12.8	14.5	16
9	8	0.3	(5,16)	6.5	8	9.3	10.4	10.5	12	14	15.5
10	7	0.2	(−20,18)	7	8	10.2	10.9	11	12.2	13.4	14.9
11	5	0.3	(−25,9)	6	7.5	9	10.4	10.5	12	14	15.5
12	8	0.2	(20,−18)	7.5	9	10	11.5	12	13.5	15	16.5
13	8	0.4	(−11,−7)	7	8.5	9.5	11	11.5	13	14	15.5
14	5	0.2	(−3,−17)	6.5	8	9.4	10.5	11	12.5	14	15.5
15	10	0.5	(2,−30)	7.5	9	11	12.5	13	14.5	16	17.5

Particle swarm optimization parameters set as: population size is 200, iterations gen = 3000, $c_1 = c_2 = 2$. There are eighteen times to achieve the average level of satisfaction in the twenty experiments, the eligible results of the experiment as shown in Table 2.

On the basis of satisfying the satisfaction level, the optimal experimental results are: two cars driving 232.35 km and total distribution cost is 1361.75 yuan, average customer satisfaction was 0.85. The optimal experimental results are shown in Table 3.

In order to compare the optimization results of the proposed model and the VRPMTW model, using the PSO algorithm to solve the case in accordance with the VRPMTW model ignored the customer fuzzy time windows, the experimental results are compared shown in Table 4. According to the comparison of the experimental results can be found that on the basis of vehicle routing problem

Table 2. Experimental results

Number of experiments	Vehicle number	Total travel distance (km)	Average customer satisfaction	Total distribution cost (Yuan)
1	2	240.83	0.87	1404.15
2	2	232.35	0.85	1361.75
3	2	232.35	0.85	1361.75
4	2	240.83	0.87	1404.15
5	2	240.83	0.87	1404.15
6	2	240.83	0.87	1404.15
7	2	232.35	0.85	1361.75
8	2	232.35	0.85	1361.75
9	2	232.35	0.85	1361.75
10	2	232.35	0.85	1361.75
11	2	239.67	0.8	1398.35
12	2	232.35	0.85	1361.75
13	2	232.35	0.85	1361.75
14	2	232.35	0.85	1361.75
15	2	239.67	0.8	1398.35
16	2	239.67	0.8	1398.35
17	2	240.83	0.87	1404.15
18	2	239.67	0.8	1398.35
Average result	2	236.33	0.84	1381.66

Table 3. Optimal experimental results

Vehicle	Optimal routings	Travel distance (km)	Cost (Yuan)
1	0-4-7-2-12-8-15-14-0	113.42	667.09
2	0-1-9-6-10-11-5-3-13-0	118.93	694.66

Table 4. Comparison of experimental results

The optimal solution of VRPMTW			The optimal solution of proposed model			
Total travel distance (km)	Total distribution cost (Yuan)	Vehicles number	Total travel distance (km)	Total distribution cost (Yuan)	Vehicles number	Average customer satisfaction
247.95	1539.75	3	232.35	1461.75	2	0.85

with multiple fuzzy time windows, keeping a higher average customer satisfaction degree can effectively reduce vehicle travel distances and lower distribution costs. Secondly, decision maker can set satisfaction degree parameters according to his own different emphasis on distribution costs and satisfaction degree and change distribution strategies. Compared to VRPMTW, this model is more flexible and realistic.

5 Conclusions

The vehicle routing problem with multiple fuzzy time windows has been widely used in the practical life, but the study of such problems are relatively few recently. Based on the multiple fuzzy time windows provided by customers, this paper adopts the membership function of start service time to quantify customer satisfaction degree, and establish universal vehicle routing problem with multiple fuzzy time windows as well as solve mathematical model by means of the particle swarm algorithm to minimize the total distribution costs and maximize customer satisfaction degree. The results of numerical example show that compared with VRPMTW, this model can reduce distribution costs more effectively. In addition, the flexible setting of customer satisfaction parameters and the expansion of the time window can further reduce the distribution cost, it has more realistic meaning and reference value for decision makers.

However, there are more complex and uncertain factors in practical applications such as vehicle travel time, the demand of customers and so on. The exploration of such extensive problems will be the main research focus on the next step.

Acknowledgements. This research was supported by NSFC (Grant No. 71401020, Grant No. 71401093) and Human Social Science for Universities of Hebei (Grant No. BJ2016057).

References

1. Ai TJ, Kachitvichyanukul V (2009) A particle swarm optimisation for vehicle routing problem with time windows. Int J Oper Res 6(19):519–537
2. Belhaiza S, Hansen P, Laporte G (2014) A hybrid variable neighborhood tabu search heuristic for the vehicle routing problem with multiple time windows. Comput Oper Res 52:269–281
3. Bent R, Van Hentenryck P (2004) A two-stage hybrid local search for the vehicle routing problem with time windows. Transp Sci 38(4):515–530
4. Bitao P, Fei W (2010) Hybrid intelligent algorithm for vehicle routing problem with multiple time windows. Int Forum Inf Technol Appl 1:181–184
5. Dantzig GB, Ramser JH (1959) The truck dispatching problem. Manage Sci 6(1):80–91
6. Desrochers M, Desrosiers J, Solomon M (1992) A new optimization algorithm for the vehicle routing problem with time windows. Oper Res 40(40):342–354

7. Doerner KF, Gronalt M et al (2008) Exact and heuristic algorithms for the vehicle routing problem with multiple interdependent time windows. Comput Oper Res 35(9):3034–3048
8. Favaretto D, Moretti E, Pellegrini P (2013) Ant colony system for a VRP with multiple time windows and multiple visits. J Interdiscip Math 10(2):263–284
9. Ghannadpour SF, Noori S et al (2014) A multi-objective dynamic vehicle routing problem with fuzzy time windows: model, solution and application. Appl Soft Comput 14(1):504–527
10. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: 1995 Proceedings of the IEEE International Conference on Neural Networks, vol 4, pp 1942–1948
11. Li Z, Zhao F, Liu H (2014) Intelligent water drops algorithm for vehicle routing problem with time windows. In: 11th International Conference on Service Systems and Service Management, vol 24, pp 1–10
12. Lin JJ (2006) Multi-objective decision making for vehicle routing problem with fuzzy due time. IEEE Int Conf Syst Man Cybern 4:2903–2908
13. Ma H, Zuo C, Yang S (2009) Modeling and solving for vehicle routing problem with multiple time windows. J Syst Eng 24(5):607–613 (in Chines)
14. Ma HW, Ye HR, Xia W (2012) Improved ant colony algorithm for solving split delivery vehicle routing problem with multiple time windows. Chin J Manage Sci 1:43–47 (in Chinese)
15. Ombuki B, Ross BJ, Hanshar F (2006) Multi-objective genetic algorithms for vehicle routing problem with time windows. Appl Intell 24(1):17–30
16. Salman A, Ahmad I, Al-Madani S (2002) Particle swarm optimization for task assignment problem. Microprocess Microsyst 26(8):363–371
17. Toklu NE, Gambardella LM, Montemanni R (2014) A multiple ant colony system for a vehicle routing problem with time windows and uncertain travel times. J Traffic Logistics Eng 2(1):52–58
18. Yan H, Gao L et al (2015) Petrol-oil and lubricants support model based on multiple time windows. J Comput Appl 35(7):2096–2100

Optimization of Operating of the Systems with Recurrent Service by Delays

Asaf Hajiyev(✉) and Narmina Abdullayeva

Institute of Control Systems, Azerbaijan National Academy of Sciences, Baku, Azerbaijan
asaf.hajiyev@gmail.com

Abstract. For systems with recurrent service delays of the beginning service are introduced. It is shown that for some systems delays reduce a customer average waiting systems. The class of so called stochastic systems for which it is advisable to introduce delays is described. Numerical and graphical examples demonstrating results are given. There are given also examples, when introducing of delays can increase a customer waiting time before service.

Keywords: Systems with recurrent service · Delays of beginning service · Customer average waiting time · Stochastic systems

1 Introduction

Consider a queuing system, where: $t_1, t_2, \cdots, t_n$ is random sequence when service starts in the system, x is customer arrival instant. At each instant ti all customers who arrived at the interval $[t_{i-1}, t_i)$ immediately will get service (see, Fig. 1). Such models are typical for applications and can be used in traffic, communication systems, network of computers and others. Denote $t_1^*, t_2^*, \cdots, t_n^*$ is (random sequence) service starts in the system with delays of beginning service (see, Fig. 1).

The problem is: Is it possible by introducing delays to reduce a customer average waiting time? May be it is a Paradoxical Idea to introduce delay of the beginning service for reducing a customer average waiting time? Idea to introduce delays of beginning service belong to various authors [1–4]. Below we will give some examples of queueing systems with delays of beginning service.

Example 1. The Main Building of M.V. Lomonosov Moscow State University. There are several lifts in the hall (see, Fig. 2).

If at least two lifts are coming to the first floor almost at the same time, then one of them must be delayed!. Mathematical model has the following form (see, Fig. 3).

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_70

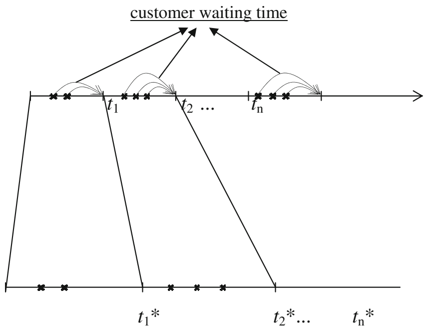


Fig. 1. General mathematical model

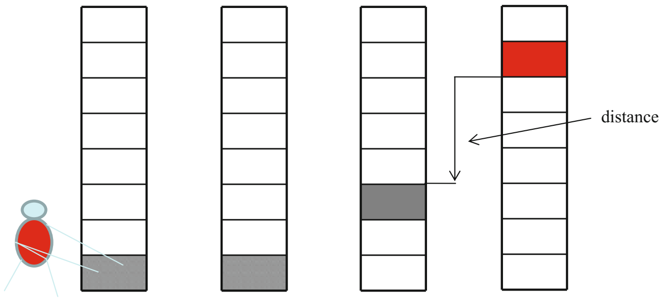


Fig. 2. Model with several elevators

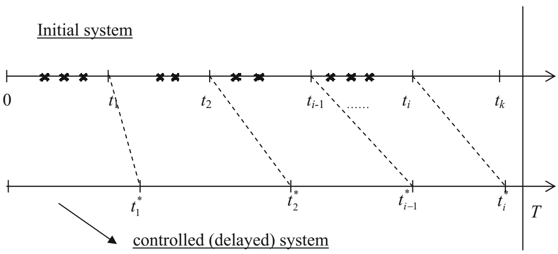


Fig. 3. Mathematical model with delays

If in the system, during the time-interval T there are k services in an initial system, but in a controlled system there are i services and $i < k$, but $w^* < w$? PARADOX? This mathematical model can be used in communication systems (Fig. 4).

2 Communication System

How long time we can keep communications?

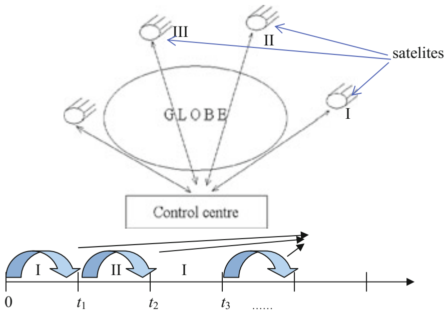


Fig. 4. Mathematical model of communication system

3 Mathematical Models

Denote $\eta_1 = t_1 \eta_2 = t_2 - t_1, \cdots, \eta_n = t_n - t_{n-1}; \eta_1, \eta_2, \cdots, \eta_n$ are independent and identically distributed random variables with distribution function $F(x)$, $E\eta_1 = \mu$, $Var\eta_1 = \delta 2$.

Stationary flow of customers arrives to service. At the each instant t_i all customers, who arrived at the interval $[t_{i-1}, t_i)$ immediately will get service. Denote w is customer average waiting time before service (Fig. 5).

From the sequence $t_1, t_2, \cdots, t_n$ we pass to the new sequence for which we have

$$\eta_1^* = t_i^* - t_{i-1}^* = \eta_i + g(\eta_i), \quad (\eta_i = t_i - t_{i-1}),$$

where $g \in G$ is a class of measurable and nonnegative functions. Denote $w(g)$ is customer average waiting time before service in a system with

$$M_F(g) = w(g) - w, \quad c = \frac{E\eta_1^2}{2E\eta_1},$$

control function $g(.)$.

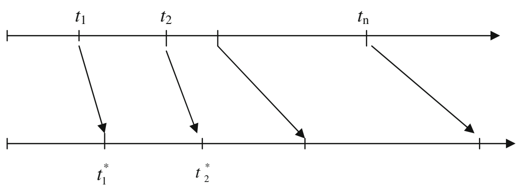


Fig. 5. Simple mathematical model with delays

Definition 1. Service can be improved if there exists $g \in G$ such that $M_F(g) < 0$.

Below will be formulated some results without proof but an explanations of these theorem will be given.

Theorem 1. *Service can be improved if and only if* $\exists x_0 < c$ *such that* $F(x_0) > 0$ *(Figs. 6 and 7).*

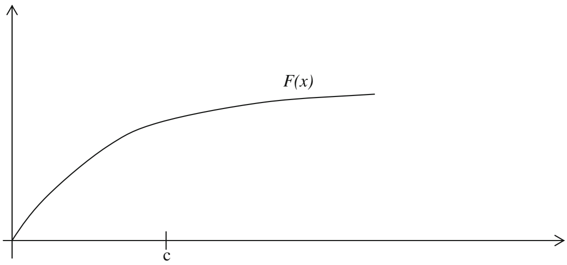


Fig. 6. Service can be improved

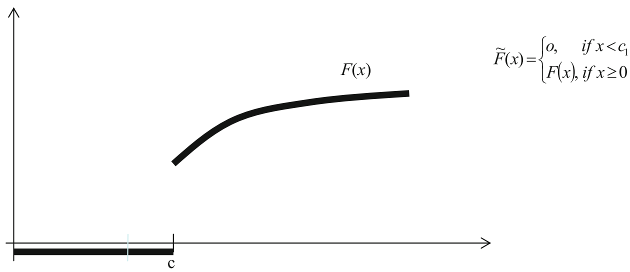


Fig. 7. Service can not be improved

Definition 2. Function $g^*(x) \in G$ is called an optimal function if

$$\min_{g \in G} M_F(g) = M_F(g^*).$$

Theorem 2. *Under the conditions of Theorem 1 an optimal function has the form* $g^*(x) = \max(0, c_1 - x) = (c_1 - x)^+$, *where* c_1 *is a unique solution of the equation*

$$c_1^2 = \int_{c_1}^{\infty} (c_1 - x)^2 dF(x).$$

Example 2. $F(x) = 1 - e - x$, $x \geq 0$. Poisson flow of customers with bounded intensity arrives to service. Simple calculations yield $w = 1$, $g^*(x) = (0,9 - x)+$, $w* = 0.9$; The gain in customer average waiting time is 10%.

Example 3. Let $t_1, t_2, \cdots, t_n, \cdots$ be times when service starts in some system. We assume $t_i = i_d^+ x_i, i = 1, 2, \cdots$, where d is a constant $x_1, x_2, \cdots, x_n$ is a sequence of independent and identically distributed random variables ($i.i.d.r.v.$) with $d.f.$

Such a scheme describes, for example, the behavior of a public transport system where $id(i = 1, 2, \cdots)$ is the ideal timetable (schedule) for the arriving transport and $id + x_i (i = 1, 2, \cdots)$ is the real timetable (schedule). We denote $\eta_I = t_i - t_{i-1} = i_d + x_i - (i-1)d - x_i - 1 = d + x_i - x_{i-1}$. Let w and σ^2 be the expectation and the variance of the customer waiting time until service. If $t_1, t_2, \cdots, t_n, \cdots$ is a stationary renewal process, $t_n = \eta_1 + \eta_2 +, \cdots + \eta_n$, then we have

$$w = \frac{E\eta^2}{2E\eta}, \ \sigma^2 = \frac{E\eta^3}{3E\eta} - \left(\frac{E\eta^2}{2E\eta}\right)^2,$$

where η is a random variable with d.f. $F(x)$. It is clear that

$$w_* = \frac{E\eta_*^2}{2E\eta_*}, \ \sigma_*^2 = \frac{E\eta_*^3}{3E\eta_*} - \left(\frac{E\eta_*^2}{2E\eta_*}\right)^2,$$

In our case $\eta_1, \eta_2, \cdots, \eta_n, \cdots$ are identically distributed random variables, but they are dependent. The flow of customers is stationary with intensity $\mu < \infty$ and independent of $t_1, t_2, \cdots, t_n$. At the instant t_i all customers, which arrived during the time interval $[t_{i-1}, t_i)$ are served immediately, in particular we assume that service time equals zero. For this case also it is succeded to get similar formula for w and σ, but we need to use for that wellknown Campell's formula.

Theorem 3. *Service in the model A can be improved by delays if and only if there exists $x_0 < c$ such, that $F(x_0) > 0$ where*

$$c = \frac{E\eta^2}{2E\eta}.$$

Definition 3. We call $\tilde{g}(x)$ an optimal function if

$$\min_{g \in G} M_F(g) = M_F(\tilde{g}).$$

Theorem 4. *Under the conditions of Theorem 1, the optimal function has the following form*

$$\tilde{g}(x) = \max\{0, \ (c_1 - x)\} = (c_1 - x)^+,$$

where c_1 is the unique solution of the equation

$$c_1^2 = \int_{c_1}^{\infty} (c_1 - x)^2 dF(x).$$

Theorem 5. *Under the conditions of Theorem 1, we have* $\sigma^2(c_1) \leq 2(c_2) \leq \sigma^2, c_2 \leq c_1$.

Example 4 (numerical). Let $t_1, t_2, \cdots, t_n, \cdots$ be times when service starts in system. We assume $t_i = id + x_i, \ \ i = 1, 2, \cdots$, where d is a constant $x_1, x_2, \cdots, x_n$ is a sequence of independent and uniformly distributed random variables on $[-1/2, 1/2]$.

Denote $\xi_i = t - t_{i-1} = (i+1)d + x_i - id - x_i - 1 = d + x_i - x_i - 1 = 1 + x_i - x_i - 1$, put $d = 1$, $E\xi_i = 1$, $\ \ E\xi_i 2 = 13/12$, $\ \ w = E\xi_i 2/2E\xi_i = 0.54$.

Simple calculations yield $g^*(x) = \max(0, 0.53 - x)$, $w^* = 0.53$; Gain in customer average waiting time is 1.8%.

3.1 Realization at the Practice 1

(1) Bus Stop

It is necessary to calculate c_1, for instance, in the Example 1, $c_1 = 0.9$; d is an interval between departure previous bus and coming to the stop next bus If $d < c_1$, then bus will be delayed for $(c_1 - d)$, otherwise nothing is changed (Fig. 8).

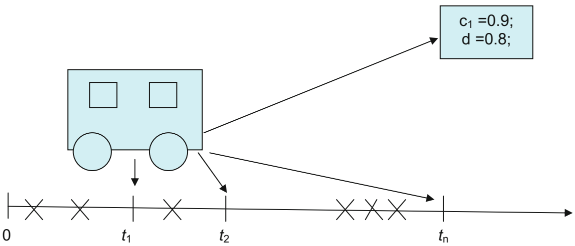


Fig. 8. Electronic plate at the bus stop

Consider two different service systems: $w_1(g)$ and $w_2(g)$ are customer average waiting time with delay function $g(x)$. Problem is: when

$$\frac{w_1(0)}{w_1(g_1^*)} > \frac{w_2(0)}{w_2(g_2^*)} \quad ?$$

i.e. where (in which system) service can be improved more strong? Introduce k, which means variation coefficient, i.e.

$$k = \frac{E\eta^2}{(E\eta)^2}; \quad \sqrt{k_1} > \frac{k_2}{1 + \sqrt{k_2 - 1}} \Rightarrow \frac{w_1(0)}{w_1(g_1^*)} > \frac{w_2(0)}{w_2(g_2^*)}.$$

Let us consider Examples 1 and 2.

Consider the random variable η has d.f. $F_1(x) = 1 - exp(-\lambda x)$, $x \geq 0$, $\lambda = 1$. $E\eta = (1/\lambda) = 1$, $E\eta^2 = 1/\lambda^2 + 1/\lambda = 2$, $k_1 = 2$. Gain in customer average waiting time was 10%. Random variable $\xi_i = 1 + x_i - x_i - 1$, where x_i has uniform distribution at the $[-1/2, 1/2]$. $E\xi_i = 1$, $E\xi_i^2 = 7/6$, $k_2 = 13/12$. Gain in a customer average waiting time was 1.5%. Our formula gives: if $(k_1)^{1/2} = 1.4 > k_2/[1 + (k_2 - 1)^{1/2}) = 0.83$; then gain is greater (Fig. 9).

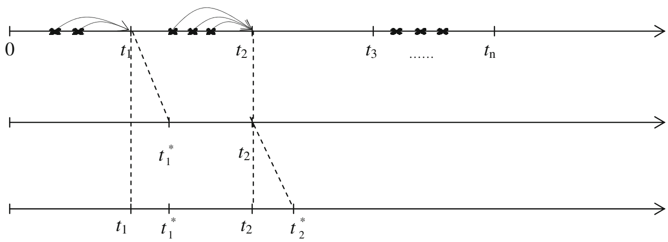


Fig. 9. Complicated mathematical model with delays

Model B:

$$\eta_1^* = \eta_1 + g_1(\eta_1),\ \eta_2^* = (\eta_1 + \eta_2 - \eta_1^*)^+ + g_2(\eta_1 + \eta_2 - \eta_1^*),$$

$$\eta_n^* = (\eta_1 + \eta_2 + \cdots + \eta_n - \eta_1^* - \eta_2^* - \cdots \eta_{n-1}^*)^+ + g_n(\_, \cdots).$$

From construction it is clear that the model has complicated form and to apply directly formula for calculation of a customer average waiting time not so easy. We will try to get new formula for calculation of a customer average waiting time before service.

Lemma 1.

$$w = \lim_{N\to\infty}\left[\frac{\frac{1}{N}\sum_{i=1}^{N} E\eta_i^2}{\frac{1}{N}2\sum_{i=1}^{N} E\eta_i}\right];$$

Generalization of Khinchin's formula

Suppose $1, 2, \cdots, n$ *is i.i.d.r.v.* $E\eta_1 =, Var\eta_1 = \sigma^2$, $F(x)$ *is d.f.*

Theorem 6. *Service can be improve* $\Leftrightarrow \sigma^2 > 0$.

Definition 4. System is called stochastic if $\sigma^2 > 0$

Remark 1. In any stochastic system service always can be improved.

Theorem 7. *For stochastic system optimal function has the form*

$$\tilde{g}(x) = (\mu - x)^+ = \max(0, \mu - x).$$

Example 5. $F(x) = 1 - e - x, \quad x \geq 0.$

$$w = 1;\ \sigma^2 = 1;\ g^*(x) = (1 - x)^+,\ w^* = 0.5;\ \sigma^{2^*} = 0.$$

The gain in customer average waiting time is 50% and variance equals zero (Fig. 10).

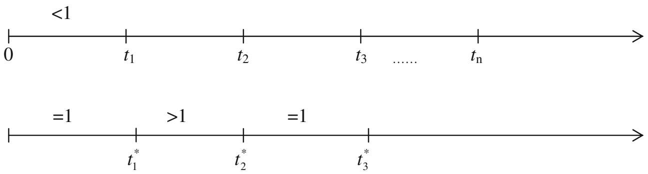


Fig. 10. Realization at the practice

(2) Shuttle Systems with Finite Volume

Denote $\eta_1 = t_1$, $\eta_2 = t_2 - t_1, \cdots$, $\eta_n = t_n - t_{n-1}$. $\eta_1, \eta_2, \cdots, \eta_n$ is independent and identically distributed random variables with distribution function $F(x)$, $E\eta_1 = \mu$, $Var\eta 1 = \delta^2$.

At the each instant ti all customers, who arrived at the interval $[t_{i-1}, t_i)$ immediately will get service, w is customer average waiting time before service .

Our aim: reduce w by introducing delay of beginning service.

Introduce delay, hence from the sequence $t_1, t_2, \cdots, t_n$ we pass to the new sequence for which we have t_1^*, t_2^*, $\cdots$, t_n^*,

$$\eta_1^* = t_i^* - t_{i-1}^* = \eta_i + g(\eta_i), \ \ (\eta_i = t_i - t_{i-1})$$

Denote m volume of Shuttle.

Theorem 8. *Service can be improved if the conditions: $\lambda E\eta_1 < m$ $x_0 < c$ such that $F(x_0) > 0$ are hold. Denote $\zeta(t)$ number of customers in a system at the instant t.*

Theorem 9. *Optimal function has the form*

$a(x) = \inf\{t : \xi(x + t) \geq m\}$ *where c_1 is a unique solution of the equation*

$$c_1^2 = \int_{c_1}^{\infty} (x - c_1)^2 dF(x).$$

3.2 Realization at the Practice 2

It is necessary to calculate c_1, for instance, in the Example 1, $c_1 = 0.9$; d is an interval between previous bus departure and coming to the stop next bus If $d < c_1$, then bus will be delayed either until the instant when number of the passengers at the interval becomes m or for $(c_1 - d)$ (Figs. 11 and 12).

Theorem 10. *Service can be improved if $\lambda E\eta_1 < m$, $\sigma^2 > 0$.*

Theorem 11. *Optimal function has the form*

$$\tilde{g}(x) = \max\{0, \min(a(x), \mu - x)\}.$$

Denote $\xi(t)$ number of customers in the system at the Instant

$$a(x) = \inf\{t : \xi(x + t) \geq m\}.$$

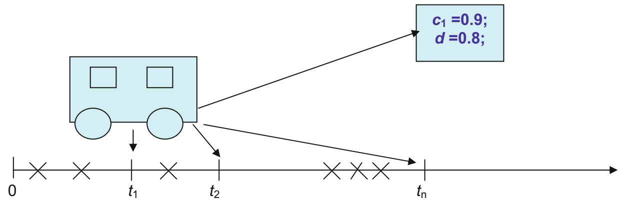


Fig. 11. Electronic plate at the bus stop

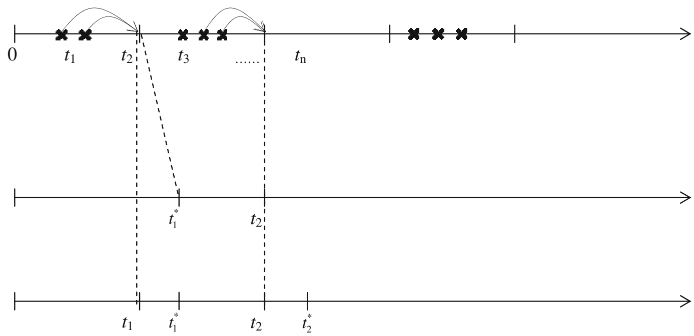


Fig. 12. Model B'

Theorem 12. *Optimal function has the form* $\tilde{g}(x) = \max\{0, \min(a(x), c_1 - x)\}$, *where* c_1 *is a unique solution of the equation* $c_1^2 = \int\limits_{c_1}^{\infty} (x - c_1)^2 dF(x)$.

3.3 Realization at the Practice 3

$E\eta = \mu$ is given or can be calculated from statistics. d is an interval between previous bus departure and coming to the stop next bus. If $d < \mu$, then bus will be delayed either until the instant when number of the passenger at the interval becomes m or for $(\mu - d)$ (Fig. 13).

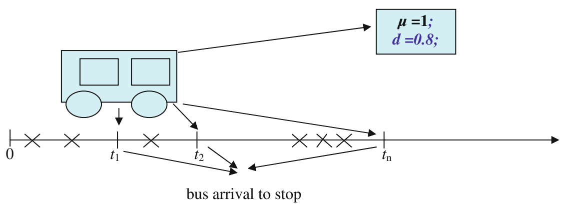


Fig. 13. Electronic plate at the bus stop

Remark 2. The proof of some results can be found in [3,4].

References

1. Carter GM, Chaiken JM, Ignall E (1972) Response areas for two emergency units. Oper Res 20(3):571–594
2. Hajiyev AH, Jafarova H (2010) Mathematical models of complex queues and their applications. Proc Inst Math Mech ANAS 33:57–66
3. Hajiyev AH, Djafarova HA, Mamedov TS (2010) Mathematical models of moving particles without overtaking. Dokl Math 81(3):395–398
4. Newell GF (1974) Control of pairing of vehicles on a public transportation route, two vehicles, one control point. Transp Sci 8(3):248–264

An Epidemic Spreading Model Based on Dynamical Network

Yi Zhang(✉)

Mathematics School, Yibin University,
Yibin 644007, People's Republic of China
245806513@qq.com

Abstract. This paper proposes a epidemic spreading model considering dynamic change of networks. A function $p(t)$ reflecting the dynamic change of network is established, which is a probability that the infected from the susceptible at time t. This function replaces the constant rate in the traditional model and a variant of the SIR model was built. This model can reflect the dynamic change of network. At the same time, a dynamic analysis of the model was conducted. Then, epidemic spreading model simulations are conducted with different parameters in ER random networks. The results show that epidemic spreads faster and more broadly in network when the smaller the parameter c, the larger parameter q. That is, epidemic often spread faster and more broadly when people take protection measures slowly and moving speed of people is slow.

Keywords: Epidemic spreading · Dynamical change · Moving speed

1 Introduction

Human beings are still at the risk of the outbreak of epidemic diseases now even after the development of modern medicine and the appearance of antibiotics [1,2]. The epidemic disease cause big disaster to human in history. For example, the SARS outbreak in 2003 and the H1N1 outbreak in 2009. These epidemic disease swept the globe in a short time, and caused a great loss, even more change the life way of human. With the worsening environment, epidemic disease is break out more frequently, and affect the humanity's survival and development. Therefore, the epidemic rule, transmission mechanism and the strategy of preventing disease are major problem which need to be solved. These problems have attracted many researchers attention increasingly.

The mathematical modeling of epidemic disease spreading has been extensively studied for a long time [13]. The main mathematical approach is the so-called compartmental models which are composed of ordinary differential equations [6–8,11]. In this kind of approach, the entire population are divided into different compartments and each compartment corresponds to an epidemiological state which depends on the characteristics of the particular disease being

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_71

modeled. The commonly used compartment; the susceptible (S), is not infected but likely infected individuals, the exposed (E) is already infection but don't have infectious individuals, the infected (I) consists of individuals who have been infected and had infectious, the removed (R) consisting of individuals who have been restored or death. A classical compartmental model is the SIR model, which was constructed by Kermack and Mckendric in 1927, and they also proposed SIS model in 1932. Based on these research, a lot of compartmental model were proposed. According to the different propagation process of disease, common epidemic model are: SI, SIS, SIR, SIRS, SEIR, SEIRS, SEI and SEIS.

However, most of research of epidemic on complex networks assume that the network structure remains the same, it does not fit the reality. Some papers have realized the dynamic nurture in epidemic spreading. A model was built to investigate the disease spreading dynamics and synchronization behavior on complex networks [12]. A SIRS epidemic model with feedback mechanism on adaptive scale-free networks is presented [9]. The real social network is the evolution of dynamic interaction between nodes and links [5]. In fact, people cut back on frequency of going out once the outbreak of epidemic, and the government take some measures to reduce mobility of people. Because of action of self-protection and government protection, the topology of the network is dynamic change. That is, the mobility of people change topology of networks, thus change the topological feature of network and the way of epidemic spreading. Therefore, it is significant that dynamic network affected by action of protection is considered in epidemic model. And it is goes some way to providing an in-depth understanding of epidemic spreading and how to control such epidemic spreading on dynamic social networks.

In this paper, We establish a function to quantify the dynamic change of the network, and propose a epidemic spreading model which consider both the change of network and the speed of the change. In this model, the function can dynamically measures the probability that the susceptible becomes the infected. A variant of the SIR model is then built, in which the traditional fixed parameters replaced with our probability function, and conduct the stability analysis of model. Then investigate how different values for parameters can effect epidemic spreading using MATLAB in ER random networks. Finally, we conclude this paper and provide several avenues for further research.

2 Epidemic Spreading Model

In this section, we construct a function of time t which considers the dynamic change of networks. Then, with this function added as a parameter to a differential dynamic spread model, a new epidemic propagation model is proposed.

2.1 Function

Consider a population of N individuals and E links respectively representing individuals and their interactions, each of individuals is in one of three states

S (susceptible), I (infected), or R (removed). At each time step t, each individual comes into contact with one other individual, drawn randomly from the population.

In epidemic spreading process, the behaviors of protection of the individuals are ignored. In reality, people will reduce the probability of going out because they are afraid of being infected when the epidemic is outbreak, and they will take some protection measures, for example, doing more exercise. With the moving of the people, the topology of network is changed, and the object they contact is also changed. Generally speaking, as time passes, because of the behavior of protection, the probability that the susceptible contact the infected is decrease gradually, and go to zero finally. Moreover, we believe that the speed at taking the measure of protection should be considered. If people do protection behavior quickly, the epidemic effect is small, but in a contrary situation, the epidemic effect is far greater.

This function is a probability $p(t)$ that an susceptible become a infected at time t, which is affected by the topology of networks. Therefore, $p(t)$ reflects the dynamical behaviors of networks at time t. In light of the above analysis, the probability at time t can be denoted as

$$p(t) = e^{-ct}, 0 < c \leq 1, t \geq 1,$$

where c is the parameter to reflect the speed of taking the measure of protection. A larger c indicates that the individual take the measure of protection in a shorter time. Figure 1 shows the probability as a function of t, given a different c. When c tends to zero, the speed of taking the measure of protection slows down. We have carefully checked that $p(t)$ decreases sharply when $c \in (0.1, 1)$, it is suggested that parameter $c < 0.1$.

In this case, the parameter for c was (0.1, 1). $p(t)$ is the probability, and in (0, 1), $p(t)$ reflects the probability that an individual become a infected from an susceptible. Therefore, $p(t)$ tends to zero as times passes.

2.2 Model

Denote $S(t)$, $I(t)$, $R(t)$ as the density of the susceptible, infected, removed at time t, $S(t)+I(t)+R(t)=1$. As shown in Fig. 2, the epidemic spreading rules can be summarized as follows:

(1) When a susceptible individual contacts an infected individual, the susceptible individual becomes a infected individual at the rate of $p(t)$, which is affected by the topology of networks. Therefore, the reduced speed of the susceptible $dS(t)/dt$ is proportional to the number of the susceptible $S(t)$ and infected $I(t)$, so we get the differential equation:

$$\frac{dS(t)}{dt} = -p(t)S(t)I(t). \tag{1}$$

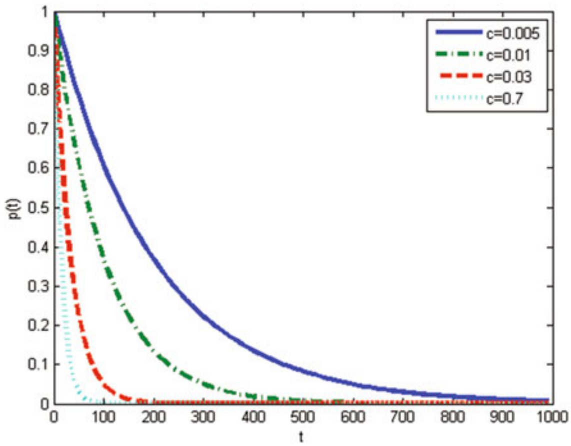


Fig. 1. Probability as a function of t

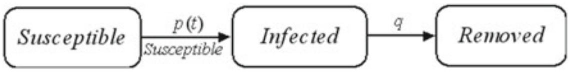


Fig. 2. Structure of epidemic spreading process

(2) At each time step, some infected individuals become removed individuals because of death or immunity. We suppose that the above case happens at probability q. Therefore, the reduced speed of the infected individuals $dI(t)/dt$ is proportional to the number of the infected individuals $I(t)$, Additionally, an susceptible individual becomes a infected individual at the rate of $p(t)$, so we have

$$\frac{dI(t)}{dt} = p(t)S(t)I(t) - qI(t). \tag{2}$$

(3) The increasing speed of the removed individuals $dR(t)/dt$ is proportional to the number of existing $I(t)$ from (2), so we get

$$\frac{dR(t)}{dt} = qI(t). \tag{3}$$

Based on the previous discussion, by integrating Eqs. (1)–(3), the following global expected model can now be formulated for the epidemic spreading process:

$$\begin{cases} \frac{dS(t)}{dt} = -p(t)S(t)I(t) \\ \frac{dI(t)}{dt} = p(t)S(t)I(t) - qI(t) \\ \frac{dR(t)}{dt} = qI(t) \\ S(0) = S_0, \quad I(0) = 1 - S_0, R(0) = 0. \end{cases} \tag{4}$$

2.3 Steady-State Analysis of Model

In the epidemic spreading process, initially the system has only the susceptible and the infected. Finally, the system has only the susceptible and the removed as it reaches an equilibrium state. So here, we analyze the system's steady-state in this period.

Because $S(t) + I(t) + R(t) = 1$, and the differential Eqs. (1) and (2) are not related to the $R(t)$, we only consider the differential Eqs. (1) and (2). Then the reduced limiting dynamic system is given by

$$\begin{cases} \frac{dS(t)}{dt} = -p(t)S(t)I(t) \\ \frac{dI(t)}{dt} = p(t)S(t)I(t) - qI(t) \\ S(0) = S_0, \quad I(0) = 1 - S_0 > 0. \end{cases} \tag{5}$$

From differential equations theory, system (4) and (5) are homogeneous, which means that analyzing the properties of system (5), is equal to analyzing the properties of system (4). First, a required theorem is introduced.

Our model is a non-autonomous differential dynamic system, the general form of which is

$$\begin{cases} \frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}) \\ \mathbf{f}(t, 0) = 0, x \in R^n. \end{cases} \tag{6}$$

Suppose that $I = [t_0, +\infty]$, $U = \{\mathbf{x} \mid \| \mathbf{x} \| \leq h\}$, and $V(t, \mathbf{x})$ is a continuous differentiable function defined in $I \times U.W(\mathbf{x})$ is a continuous differentiable function defined in U.

Theorem 1. [10] *Suppose that $V(t, \boldsymbol{x})$ is a positive definite function in $I \times R^n$, which has an infinitesimally small upper bound and an infinite lower bound, and $\dot{V}(t, \boldsymbol{x})$ is a negative semi-definite, so the zero solutions for system (6) are global uniformly asymptotically stable.*

According to the differential equations theory, we can get two theorem as follows.

Theorem 2. *$P^* = (S^*, 0)(0 \leq S^* < 1)$is the equilibrium of the system (5).*

Proof. Let the right side of each of the differential Eq. (8) be equal to zero in the system which gives the equation

$$-p(t)S(t)I(t) = 0, \tag{7}$$

$$p(t)I(t) - qI(t) = 0. \tag{8}$$

The feasible region for the equations is R^2, and so we study the equations in a closed set $A = \{(S, E, I) \in R^2 \mid S + I \leq 1, S, I \geq 0\}$

From Eq. (8), we can get $I(t) = 0 (p(t) \neq q)$. Substituting $I(t) = 0$ into Eq. (7), we can get the system has the equilibrium $P^* = (S,\ I) = (S^*, 0)$ $(0 \leq S^* < 1)$. □

Therefore, we can obtain the equilibrium for the system (4) $Q^* = (S^*, 0, R^*)$, where $S^* + R^* = 1$; that is, the epidemic must disappear with time, and all S^* and R^* that satisfy $S^* + R^* = 1$, represent stable situations. Next, we look at the stability of the equilibrium $P^* = (S^*, 0, 0)$.

Theorem 3. *The equilibrium P^* is globally uniformly asymptotically stable.*

Proof. Let the Liapunov function be $V(S, I) = F(S) + G(I)$, and taking the derivative of V versus t along the solution for the equations, we have

$$V'(S, I) = -F'(S)p(t)S(t)I(t) + G'(I)[p(t)S(t)I(t) - qI(t)].$$

In order to ascertain $V(S, I)$ and $V'(S, I)$, and taking the value $F'(S) = 1$, $G'(I) = 1$, we get $V(S, I) = S + I$, and

$$V'(S, I) = -p(t)S(t)I(t) + p(t)S(t)I(t) - qI(t) = -qI(t) \leq 0.$$

So we get $V(S, I)$, which is positive definite, $0 \leq V(S, I) \leq 1 + S$, that is, it has an infinitesimally small upper bound and an infinite lower bound, and $V'(S, I)$ is a negative semi-definite in the feasible region $A = \{(S, I) \in R^2 | S+I \leq 1, S, I \geq 0\}$.

From Theorem 1, the equilibrium P^* is globally uniformly asymptotically stable. □

3 Results and Analysis

The probability that a susceptible individual becomes a infected individual is given a constant over the whole spreading process in traditional epidemic propagation models, but dynamical network needs to be considered in the model as

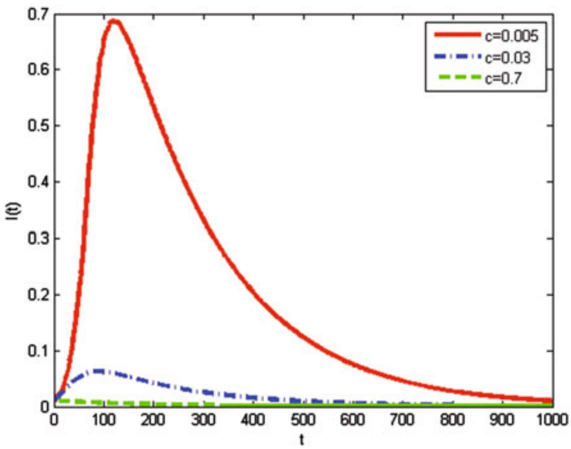


Fig. 3. Density of the infected with different c on a random network

can be seen in the above analysis. Therefore, the probability is a variable which changes over time in our model. In the following sections, we examine the results with a different values for parameter c to demonstrate how the act of isolated affects epidemic propagation. Simulations were carried out using MATLAB.

To compare the results, the proposed model with different value for c was tested on a artificial networks–ER random networks with the network size at $N = 10000$ and the average degree $<k> = 16$ [4]. Two different nodes are connected with a probability of $p = 0.0016$, so we can get random networks with N nodes, and $\frac{pN(N-1)}{2}$ edges. ER random networks degree-distributions are an approximate Poisson distribution $<k> = pN \approx p(N-1)$. This distribution reaches its peak value at the average degree $<k>$ [3].

Given that the other parameters are fixed, we compared the epidemic spreading processes on the random networks with different values for c. So here, we set $c = 0.005$, $c = 0.03$, $c = 0.7$ when $q = 0.05$ respectively. Figure 3 illustrates how the density of the infected changes over time for different values of c in a random network. From a macroscopic perspective, we found that as parameter c decreased, the number of the infected peaked increased, because a smaller c indicates a faster movement speed. The green line indicates a scenario in which the individuals move quickly, that is, the topology of network change quickly. The blue line represents a scenario in which the individuals move speed is slow relatively, and the red line represents a scenario in which the individuals hardly move. It can be seen that the higher parameter c is, the smaller the infected peak value. Figure 4 describes how the density of the removed changes with changes in the parameter c. We found that as parameter c decreased, the number of the end the size of the removed increased, because a smaller c indicates a faster movement speed. The final value for the removed density $R(t)$ is greater, which indicates the number of people affected by an epidemic when c is smaller.

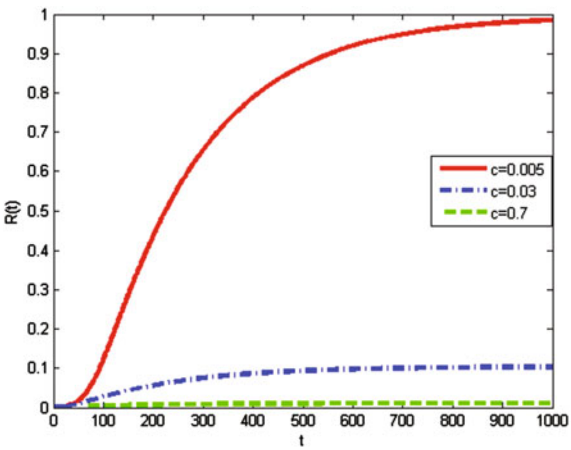


Fig. 4. Density of the removed with different c on a random network

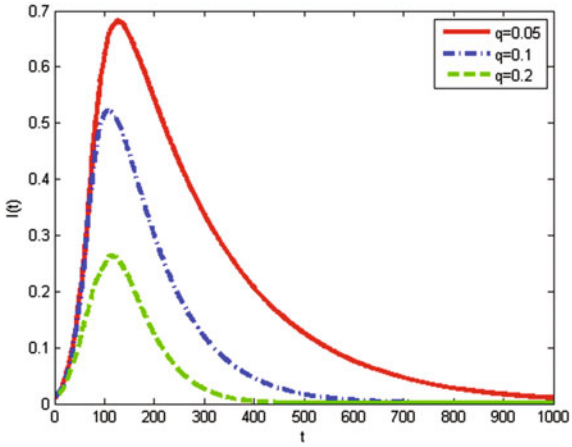


Fig. 5. Density of the infected with different q on a random network

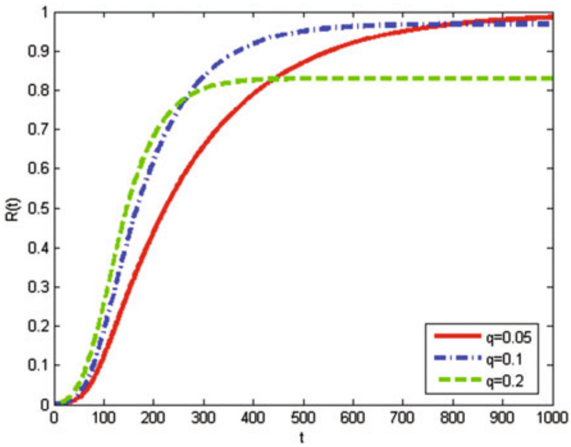


Fig. 6. Density of the removed with different q on a random network

Clearly, the smaller the value of c when other parameters are fixed, the broader the epidemic's influence. A smaller c indicates that if the individuals have a faster movement speed, less susceptible individuals change into removed individuals, and therefore the influence of the epidemic decreases. Generally speaking, it can be seen that epidemic spread more broadly and last longer when c is smaller. However, there is a more significant impact from protection measures in random networks.

Given that the other parameters are fixed, we compared the epidemic spreading processes with different values for q. we set $q = 0.05$, $q = 0.1$, and $q = 0.2$. Figures 5 and 6 illustrates how the density of the infected and removed changes over time for the different values of q. From the Fig. 5, we can see that the bigger the value of q, when other parameters are fixed, the smaller peak value of the infected. It can be seen that the higher parameter q is, the larger the infected peak value, and the slower the epidemic terminates. From the Fig. 6, it is easy to find that the final value for the removed density $R(t)$ is greater, which indicates the number of people is infected when q is smaller. Clearly, q is the transformation probability that infected individuals become removed individuals, in reality, along with the increase of q, there is more infected individuals become removed individuals. As a result, the number of removed increases, the influence of the epidemic decreases.

4 Conclusions

In this paper, we proposed a variant of the SIR epidemic spreading model by considering the dynamic change of networks because of the behavior of protection of individuals. In our model, we established a function $p(t)$ to describe the dynamic change of networks. $p(t)$ also was a probability that the infected from the susceptible at time t. The parameter c of $p(t)$ reflecting the change speed of taking measure of protection. If the individuals take protection measures quickly, c is greater, with the epidemic effect is small. On the contrary, if c is smaller, the epidemic effect is greater. Generally speaking, as time passes, because of the behavior of protection, the probability that the susceptible contact the infected is decrease gradually, and go to zero finally. This function as a parameter was added to dynamic differential equations, and a variant of the SIR model was built. At the same time, a dynamic analysis of the model was conducted.

We then do the simulation results from our epidemic spreading model using a different values for c, q in ER random networks. The simulations showed that epidemic spread faster and more broadly when c is smaller. That is, epidemic often spread faster and more broadly when people take protection measures slowly. At the same time, the simulation results suggested that the spread of epidemics is influenced significantly by parameter c in a random network. q is the transformation probability that infected individuals become removed individuals, in reality, along with the increase of q, the infected peak value is large, and the epidemic terminates slowly. These results indicated that epidemic spreads faster and more broadly in network when the smaller the parameter c, the larger parameter q.

The dynamic networks plays an important role in epidemic spreading. With this in mind, in the future we plan to study feature of dynamical networks. In addition, further studies will be conducted on epidemic control strategies according to different network topologies.

Acknowledgements. This research was supported by project of Education Department in Sichan province (Grant No. 15ZB0295). We appreciated all support in finance and in spirit.

References

1. Albert R, Barabási AL (2002) Statistical mechanics of complex networks. Rev Mod Phys 74(1):47–97

2. Boccaletti S, Latora V, Moreno Y et al (2006) Complex networks: structure and dynamics. Phys Rep 424(4–5):175–308
3. Bollobas B (2001) Random Graphs. Academic Press, New York
4. Erdös P, Rényi A (1959) On random graphs I. Publicationes Math 6:290–297
5. Gao L, Liu JM, Zhang SW et al (2010) Dynamic changes of social network structure caused by memory characteristics. Comput Eng 36(10):11–13
6. Hethcote HW (2000) The mathematics of infectious diseases. Siam Rev 42(4):599–653
7. Korobeinikov A, Wake GC (2002) Lyapunov functions and global stability for sir, sirs, and sis epidemiological models. Appl Math Lett 15(8):955–960
8. Li CH, Tsai CC, Yang SY (2014) Analysis of epidemic spreading of an sirs model in complex heterogeneous networks. Commun Nonlinear Sci Numer Simul 19(4):1042–1054
9. Li T, Liu X, Wu J, Wan C et al (2016) An epidemic spreading model on adaptive scale-free networks with feedback mechanism. Phys A Stat Mech Appl 450:649–656
10. Ma ZE, Zhou YC (2001) The Qualitative and Stability Theory of Ordinary Differential Equations. Science press in Chinese
11. Moneim IA (2007) Seasonally varying epidemics with and without latent period: a comparative simulation study. Math Med Biol 24(1):1–15 A Journal of the IMA
12. Xu D, Xu X, Yang C et al (2016) Spreading dynamics and synchronization behavior of periodic diseases on complex networks. Phys A Stat Mech Appl 466:544–551
13. Zhu G, Fu X, Chen G (2012) Spreading dynamics and global stability of a generalized epidemic model on complex heterogeneous networks. Appl Math Model 36(12):5808–5817

Analysis of Enterprise Microblog Marketing in Different Industries Based on DEA Model

Dan Zhang(✉), Yufeng Ma, Aixin Wang, Yue He, and Changzheng He

Business School, Sichuan University,
Chengdu 610065, People's Republic of China
253510764@qq.com

Abstract. In this paper, we propose a evaluation system for the enterprise microblog marketing effectiveness, combine DEA method to establish the evaluation model, and applying the quantitative evaluation method of input and output to evaluate the effect of microblog marketing among different enterprises. In the empirical part, in order to verify the feasibility of the model, we compare the effectiveness of the relative marketing effect of 50 official microblog in 5 different industries based on Sina microblog. The result shows that microblog marketing effect among different enterprises is different. In the five major industries, the integrated marketing effect of mobile phones industry is best, followed by Home Furnishing, skin care, snacks, the worst is the women's clothing industry. In the end, by combining the development characteristics and situation of each industry, the paper puts forward some measures to improve the marketing effect of enterprise microblog.

Keywords: Microblog marketing · Enterprise microblog · Marketing effectiveness · DEA

1 Introduction

CNNIC [10] has pointed out that as of June 2016, Chinese netizens has reached 710 million, the size of microblog users to 242 million. The huge user resources of microblog provide a rich fertile soil of marketing activities for enterprises. Many companies have realized the importance of microblog. Therefore, they launch the online marketing on microblog. However, how to assess the effectiveness of microblog marketing becomes the practical problems faced by enterprises.

With the Matthew effect of microblog is gradually prominent, the number of the domestic and foreign scholars who study on the effect of microblog marketing are on the rise. Scholars on microblog marketing research mainly focused on three aspects: marketing effect, the impact of microblog marketing and marketing strategy. In the study of the marketing effect of microblog, Leung [13] attempted to explore the marketing effectiveness of two different social media sites (Facebook and Twitter) in the hotel industry, Integrated the attitude-toward-the-ad

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_72

(Aad) model with the concepts of attitude-toward-social-media-page, and finally proposed a theoretical model of hotel social media marketing effectiveness. For different groups of people, microblog marketing may have different influence, Barry [2] assessed whether alcohol companies restrict youth/adolescent access, interaction, and exposured to their marketing on Twitter and Instagram. Clark [9] found that due to the youth presence on Twitter and the clinical uncertainty of the long term health complications of electronic cigarette consumption, the protection of public health warrants scrutiny and potential regulation of social media marketing. Chen [8] explored young consumers' interpretation of Twitter and marketing information on this particular social media platform from the perspective of consumers. Good marketing strategy is very important for enterprises, Liu [14] developed a guideline for advertisers to make the best marketing practices by visualizing and studying the evolution of popular Twitter hashtags. In order to understand how social media affect the hospitality and tourism field has increased, Park [15] discussed and demonstrated social media analytics using Twitter data referring to cruise travel, finally provided feasible marketing strategies. Kafeza [12] introduced a novel methodology to achieve information diffusion within a social graph that activates a realistic number of users, the methodology is useful to marketers who are interested to use social influence and run effective marketing campaigns.

To sum up, the innovation of this paper is to put forward the microblog marketing activity is input and output activities, through the introduction of DEA model, we establishes the comprehensive evaluation index system using quantitative method to calculate the input-output efficiency of enterprises microblog marketing, and then evaluate the effect of enterprise microblog marketing.

2 Related Work

2.1 Enterprise Microblog Marketing Effect Evaluation System

There are relatively abundant indicators to assess the effect of microblog marketing. Through quantities of case studies and quantitative research methods, some scholars found that there are many factors which have influence on the enterprise microblog marketing effectiveness. Xue [18] found that sales promotion is easy to promote consumer to make information share by forwarding to a friend or to make comments. In addition, Chang [5] thought liking or sharing social media messages can increase the effects of popular cohesion and message diffusion. What's more, Zhu [20] argued that social media marketing efforts need to be congruent and aligned with the different needs of social media users.

Bi [3] thought that the comments and reposts can reflect the microblog marketing effects. Saulles [16] considered not only the simply counting Twitter followers and volumes of tweets as indicators of effectiveness but also utilises social authority scoring from the digital marketing analysts. Based on the above findings, this paper selects the number of fans, the number of microblog, the frequency of releasing microblog, and prizes value as the input evaluation indexes. The selected output evaluation indexes are the number of fans forwarding, the number of fans comments, the number of fans point of praise, and non negative sentiment index.

From the input index perspective, this paper puts forward active comments and passive comments. Active comments means that the numbers of blogger's active replies to followers. Here the follower's comments do not express the tendency of doubt or the desire to get any responds from the blogger. However, the passive comments refer to the number of the replies. The replies are the blogger who make to answer the doubts or the questions which fans put forward to. In fact, there are obvious marks in these comments. For example, these comments consist of the symbol@, or the interrogative and obvious punctuation, microblog expressions.

From the output index perspective, the original microblog forwarded and the original microblog comments were proposed. The reason is that the official microblog generally contain original and non-original microblog(namely forwarding other's microblog). The total numbers of the original microblog forwarded also contain the original microblog forwarded and the non-original forwarded. Therefore, in order to more clearly measure the spread of the original microblog, the article puts forward the two indicators.

Based on the above analysis, the evaluation index system of enterprise microblog marketing effectiveness is established, as is shown in Fig. 1.

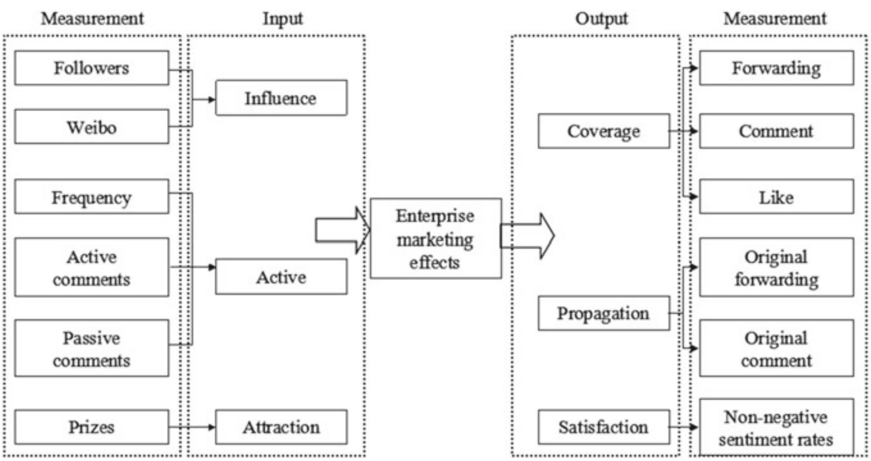


Fig. 1. Enterprise microblog marketing effect evaluation index system

2.2 Analytic Approach

Data Envelopment Analysis which can be used for evaluating the relative efficiency of decision-making units in organizations was first proposed by Charnes et al. [6]. While DEA prototype can be traced back to 1957, in the analysis of British agricultural productivity, Farrell [4] put forward the thought of envelope.

Later, the related DEA methods are mostly building on the idea of Farrell to develop a mathematical programming model. Other than comparing efficiency across DMUs within an organization, DEA has also been used to compare efficiency across firms.

DEA model has many forms. This paper selects the BCC model which can evaluate the effectiveness of scale and technical. Sun [17] proposed that we can suppose there are n decision-making units, each unit $\mathrm{DMU}_j(j = 1, 2, \cdots, n)$ has m inputs and s outputs. Respectively, they are showed by input X and output Y.

$$X_j = (x_{1j}, x_{2j}, \cdots, x_{nj}), \quad Y_j = (y_{1j}, y_{2j}, \cdots, y_{nj})^T, \quad j = 1, 2, \cdots, n.$$

To evaluate DMU_j, Cooper [11] thought that we can use the linear programming model (P) and dual programming model (D), as shown below:

$$(p) = \begin{cases} \max \mu^T y_0 = V_p \\ \omega^T x_j - \mu_i^T \geq 0 (j = 1, 2, ..., n) \\ \omega^T x_0 = 1 \\ \omega \geq 0, \mu \geq 0, \end{cases} \tag{1}$$

$$(D) = \begin{cases} \min V_D = \theta \\ \sum\limits_{j=1}^{n} x_j \lambda_j + s^- = \theta x_0 \\ \sum\limits_{j=1}^{n} y_j \lambda_j - s^+ = y_0 \\ \lambda_j \geq 0, (1 \leq j \leq n) \\ s^+ \geq 0, s^- \geq 0. \end{cases} \tag{2}$$

In the above formula, s^- and s^+ are slack variable. λ_j is the weights coefficient of input and output indicators. θ represents the ratio of input reduction. If $\theta^* = 1, s^{-*} = s^{+*} = 0$, the unit is considered to be valid. If $\theta^* = 1$, the s^{-*}, s^{+*} are not all 0, the unit is considered weak effective. If $\theta^* < 1$, the unit is considered to be invalid. When using the model to measure technical efficiency, technical efficiency can be decomposed into pure technical efficiency and scale efficiency.

3 Data Collection

3.1 Sample Selection

The research time was setting on March 5, 2014-March 11, 2014. When choosing the brands, this paper discretionarily chooses five different industries, which is Women's Clothing, Mobile Phones, Skin Care, Home Furnishings, and Snacks. Then the microblog of enterprises who have the sina official certification will be determined to as the study sample. The final choice of enterprises list is as shown in Table 1.

Table 1. The chosen enterprises in this research

Category	Women's clothing	Mobile phones	Skin care	Home furnishings	Snacks
1	Ochirly	Apple	EsteeLauder	Lin shi mu ye	Three squirrel
2	HSTYLE	Samsung	Pechoin	Bell Land	Be & Cheery
3	Vero Moda	MIUI	Lancome	Gudxon	Haoxiangni
4	ELF SACK	Huawei	Laneige	Coleshome	FERRERO ROCHER
5	Artka	VIVO	L'Oreal	QuanU	ZHOUHEI YA
6	INMAN	Coolpad	Kiehl's	Flisa	Lou lan mi yu
7	GIRDEAR	Nokia	MEIFUBAO	Hegou	BESTORE
8	ONLY	HTC	Marykay	KUKa	Xinnongge
9	LIEBO	Sony	AFu	LUHU	Xiyumeinong
10	PeaceBird	Lenovo	CHCEDO	YIMILOVE	Houstage

3.2 The Experimental Data Preprocessing

This paper collects 1363 microblog and 128483 comments on the 50 sina official microblog. In the all comments, there are 123782 fans comments, 2584 active comments and 2117 passive comments. The definition of main input and output indicators are as follows:

(1) The input indicators
Followers and microblog are all gotten from the pameng crawler.

The frequency of releasing microblog (FRW). This paper collects microblog of enterprise during 7 days. The frequency of post microblog = the numbers of microblog/7.

The number of active comments (NAC). This paper invites two researchers to judge the emotion of the comments. When judge the comments, the researcher according to the emotion and tone in the sentence pretend to be a follower to determine whether the comment has the tendency to question.

The number of passive comments (NPC). Use the similar collecting method with the active comments.

Prizes value (PV). Through sina advanced search, screening the microblog which contain prizes information. Then, recording the number of prizes and price. At last, this paper calculates the price of the total prize according to the market price.

(2) The output indicators
The number of microblog forward (NWF), the number of fans praise (NWL), the original microblog forwarded (OWF), the original microblog comment (OWC). They are all gotten from the pameng crawler.

The number of follower comments (NWC). Follower comments = microblog comments − active comments − passive comments.

Non negative sentiment rates (NNS). Non negative sentiment rates = positive emotion + neutral emotion)/all kinds of emotion. This represents the ratio of non-negative sentiment occupying the total emotional. This rate can also reflect the ratio of negative emotion from the side. The enterprise collects negative emotions to effectively understand the user, and then help enterprises to improve the service.

In order to improve the accuracy of the emotional data, this paper use artificial method to judge the emotion in the comments. When collecting the comments, this paper regards the spam comments as the neutral emotion to avoid its interference.

After the comments collection is to classify the emotion of the comments. This paper use the content analysis method. Then, these two researchers will classify the comments which belong to one enterprise that is chosen randomly among 50 enterprises. These comments will be used to practice the researcher repeatedly to ensure the accuracy. Then, the Kappa Statistic in reliability studies will be used to test the researcher. As a result, the researcher's kappa value is 0.87. That shows a good consistency. Therefore, these two researchers can be used in the phase of emotional classification. Finally, we can get the number of positive comments, negative comments and neutral comments. Average these three comments are the brand's corresponding comments. We can use these data in our study.

Then, each brand's microblog indexes in 7 days (except for the frequency to send microblog, the prize value, non-negative sentiment index) are also on average. Due to the numerous tables, this paper only selects the ladies enterprises' data, as shown in Tables 2 and 3.

From the table we can see, there are many data whose value is 0 in the column of active comments, passive comments and prize value. That is because that different enterprise has different marketing strategy. Thus, some input data value of enterprises is 0. Besides, this paper uses the infinitesimal (0.0001) to replace the 0 due to the forbiddance that the input DEA data cannot be 0.

3.3 Experimental Result

This paper uses the transverse, longitudinal comparison to do the comprehensive analysis in the part of results analysis.

Table 2. The input indicators of Women's Clothing enterprises

	Followers (Ten thousand)	Microblog (Thousand)	FRW	NAC	NPC	PV (Hundred)
Ochirly 11.96 2.05	0.71	1.6	0.2	0		
HSTYLE	14.67	0.26	9.29	0.42	0.11	3.65
1Vero Moda	15.61	2.52	1.71	0.08	0	
ELF SACK	5.55	9.82	8.86	2.89	0.27	2.5
Artka	3.7	2.11	7.14	0.34	0.04	14.46
INMAN	22.26	6.33	8.71	1.92	0.08 5	
GIRDEAR	1.05	3.91	0.29	0	0	0
ONLY	47.26	4.56	5	0.11	0	0
LIEBO	27.52	6.72	2.71	0.37	0.21 9	
PeaceBird	6.86	4.01	2	0	0	0

Transverse analysis is a method of comparative analysis of different enterprises' microblog marketing effect between input and output efficiency within the same industry. Longitudinal analysis refers to the comparative analysis of different industries. First, arranging all the 50 brands DEA results data in a table. Then regarding the comprehensive technical efficiency as the key word, according to the descending order, to compare the input and output efficiency of different industries.

(1) The Transverse Comparison
This paper is based on BCC model of DEA which assumes that scale reward is variable. DEAP2.1 software is applied to this experiment. Through calculating, we can obtain the evaluation results. Because of the 50 brands, the number of relate tables are also very large. In this thesis, the representative list is only shown in Table 4.

Table 3. The output indicators of Women's Clothing enterprises

	NWL	NWF	NWC	OWF	OWC	NNS
Ochirly	9.2	6.4	11.2	7.75	13	0.93
HSTYLE	3.22	10.38	4.14	14.5	5.64	0.99
Vero Moda	4.42	3.58	4.33	3.58	4.33	0.63
ELF SACK	3.73	3.95	13.9	4.68	15.3	0.99
Artka	3.5	3.64	3.52	3.36	3.06	0.98
INMAN	7.36	125.56	35.69	84.4	44.09	0.92
GIRDEAR	0.5	0.5	2.5	0.5	2.5	1
ONLY	6.97	10.2	3.63	10.47	3.74	0.91
LIEBO	4	17.16	15.68	43.57	39.86	0.99
PeaceBird	0.29	0.29	0.43	0.29	0.43	0.5

Table 4. The marketing efficiency of Women's Clothing

Enterprises	Crste	Vrste	Scale	
Ochirly	1	1	1	-
HSTYLE	0.895	0.898	0.997	irs
Vero Moda	1	1	1	-
ELF SACK	0.522	0.522	1	-
Artka	0.543	0.545	0.997	irs
INMAN	0.45	0.453	0.994	irs
GIRDEAR	1	1	1	-
ONLY	1	1	1	-
LIEBO	0.31	0.31	0.998	irs
PeaceBird	0.5	1	0.5	irs

If the value of the comprehensive efficiency of an enterprise is 1, that shows that the input-output efficiency of the enterprise marketing activities is ideal, namely the enterprise marketing effect is good. The irs in the table represents the enterprises marketing activity is in the state of increasing, drs represents the decreasing state.

The comprehensive technical efficiency is the measurement and evaluation of the ability of resource allocation and resource utilization of the decision-making unit. Pure technical efficiency is the production efficiency of enterprises due to the factors such as management and technology, and the specific meaning is the ability to get the maximum output under the given input. Scale efficiency is the production efficiency which is influenced by the scale of enterprise. It reflects the gap between the actual size and the optimal production scale. According to the production status of activities in the part of returns to scale, the scale efficiency can be divided into three states, increasing returns to scale (irs) and decreasing returns to scale (drs) and constant returns to scale (crs).

It can be seen from Tables 3 and 4, in the 10 brands of Women's Clothing category, only 4 enterprises reach the ideal input-output ratio. They are Ochirly, VEROMODA, GIRDEAR and ONLY. HSTYLE, Artka, INMAN, LIEBO's marketing effect are not ideal, because the pure technical and scale efficiency are invalid. In addition, they are in a state of increasing return of scale. This shows that its allocation of resources and the scale of investment are not good enough. The low comprehensive efficiency of ELF SACK is due to the invalidation of scale. The best method to improve the condition is to increases the scale of activities and expands its exposure rate.

After the comprehensive comparison of the data table of other industry, the article finds that in five industries, the integrate marketing effect of mobile phone is best. 80% enterprises achieve the satisfied efficiency. Then the following is Home Furnishing 70%, Skin Care 60%, and Snacks 50%. The worst result is Women's Clothing industry, only 40%. The low satisfaction of Women's clothing industry may be due to the fans' activity participation is not high because of its boring and unattractive contents. Its microblog mostly show new products, seldom interacting with fans. So the official microblog is similar to enterprise's advertising platform, rather than a communication platform with fans. Lacking of interaction and humanization, the marketing effect of the product is naturely not very good.

(2) The Longitudinal Comparison

In order to better evaluate the microblog marketing effect between the five major industries, this paper proposes to use the overall ranking of comprehensive technical efficiency in DEA results in the table to identify the success or failure of enterprise marketing effect. Ranking the Comprehensive technical efficiency value of the 50 brands from high to low sequence, if the technical efficiency is equal, the brands are in same ranking. After that, the ranking will be as the enterprise scores. Higher ranking namely that the ranking order is more backward explains that the enterprise marketing effect is poorer. Calculate every

Table 5. The ranking status of comprehensive technical efficiency which is not 1

Ranking	Enterprises	Crste	Vrste	Scale	
31	Pechoin	0.99	1	0.99	irs
32	Lin shi mu ye	0.968	1	0.968	irs
33	Haoxiangni	0.956	1	0.956	irs
34	Lenovo	0.955	1	0.955	irs
35	HSTYLE	0.895	0.898	0.997	irs
36	Nokia	0.673	0.677	0.994	drs
37	Laneige	0.655	0.704	0.93	drs
38	EsteeLauder	0.589	0.591	0.998	drs
39	Artka	0.543	0.545	0.997	irs
40	ELF SACK	0.522	0.522	1	-
41	PeaceBird	0.5	1	0.5	irs
42	INMAN	0.45	0.453	0.994	irs
43	MEIFUBAO	0.444	0.445	0.999	-
44	Three squirrel	0.414	0.414	0.999	irs
45	LIEBO	0.31	0.31	0.998	irs
46	BESTORE	0.243	0.244	0.996	drs
47	Be & Cheery	0.142	0.143	0.999	-
48	Xinnongge	0.127	0.128	0.997	-
49	Bell Land	0.021	1	0.021	irs
50	Gudxon	0.001	1	0.001	irs

enterprise's score, and then classified calculate the score of each category. The higher score shows that the category's marketing effect is poorer.

Experimental data shows that the top 30 enterprises' marketing effects are very ideal, because the comprehensive technical efficiency, pure technical efficiency and scale efficiency value are 1. The comprehensive technical efficiency of last 20 enterprises are different, as shown in Table 5.

The data shows that in the top 30 enterprises, mobile phones companies have 8, accounting for 26.7%; Home Furnishing class has 7, accounting for 23.3%; skin care companies have 6, occupying 20%; snacks has 5, accounting for 16.7%; women's clothing enterprises have 4, occupying 13.3%.

In whole 50 enterprises, achieving the satisfactory efficiency of enterprises in the mobile phones industry accounted for 16%, followed by home furnishing for 14%, skin care for 12%, snacks for 10%. Women's clothing class accounts for only 8%. The above data shows that in five industries, the microblog marketing effect of mobile phones companies are doing relatively well, women's relatively weak.

Then, classifying the ranking scores of all brands, mobile phone has the lowest score 78 points. Home furnishing gets 138 points. Skin care is 155 points. Snack is 223 points. Women's clothing gets the highest score 246 points. Therefore, from the perspective of input and output, the input-output ratio of mobile phones is the most ideal, which show that there are good microblog marketing effect in this category. The home furnishing and skin care industries are also relatively well. The input of snacks, women's clothing is redundant, while the output is insufficient. They should properly adjust the resources reasonably, display the resources value to maximize to increase enterprise microblog marketing effect.

In summary, this paper finds the marketing effect of different goods of different types of enterprises on sina microblog is different. The best effect is mobile phones industry, followed by Home Furnishing, skin care, snacks. The worst is the women's clothing industry.

3.4 Results

(1) Microblog Marketing Input-Output Analysis
As can be seen from the previous section, the invalid marketing effects of enterprises have 3 kinds:

- The technical efficiency, pure technical efficiency, scale efficiency are all not 1;
- The technical efficiency, pure technical efficiency is not 1, scale efficiency is 1;
- The pure technical efficiency is 1, the technical efficiency, scale efficiency is not 1.

The paper selects the first representative case to analyze. There are some enterprises satisfying the condition, randomly select one company. As a result, this paper chooses Be & Cheery as the sample. Table 6 is Be & Cheery's unit adjusted value: Table 6 The adjustment values of Be & Cheery.

The comprehensive efficiency value of Be & Cheery is 0.142. The pure technical efficiency value is 0.143. The scale efficiency value is 0.999. The comprehensive efficiency and pure technical efficiency value are very low, that shows there is a big problem. From the table we can see, there are input redundancy in 6 input indexes, especially on the number of followers. The original value of followers is 80.650, but the actual target value is 5.508. Be & Cheery spends too much energy on the increasing number of followers, and this part of input do not actually convert into the corresponding output, which cause a great waste of resources. Be & Cheery releases 42 microblog within 7 days, with high number of active and passive comments. Relatively speaking, official microblog is active, but the active behavior does not bring good interaction. The average number of forwarding and comments are all only 5, the number of praise is 2. The output is significantly inadequate. Besides, the prizes of Be & Cheery are attractive, the price reaching 1489 yuan. However, they do not get corresponding fans participation. In addition, Be & Cheery is good at guiding the followers' emotion and controlling the scale of marketing. The future improvement is mainly on the allocation of resources. To sum up, the above adjustment values can provide an important basis for the evaluation of enterprise marketing effects.

Table 6. The adjustment values of Be & Cheery

Enterprises	Indicators	Original value	Redundant input value	Insufficient output value	Effective target value
Be & Cheery	NWL	1.881	0	0.965	2.846
	NWF	5.929	0	0.005	5.934
	NWC	5.714	0	0.953	6.668
	OWF	6.639	0	0.015	6.654
	OWC	6.25	0	0	6.625
	NNS	0.992	0	0	0.992
	Followers	80.65	−69.15	−5.992	5.508
	microblog	6.13	−5.256	0	0.874
	FRW	6	−5.144	0	0.856
	NAC	0.667	−0.572	0	0.095
	NPC	0.262	−0.225	0	0.037
	PV	14.89	−12.767	−1.563	0.56

(2) The Improve Measurement for Enterprises
Through calculation of the DEA model, this paper finds that the microblog marketing effect of firms in different industries is different. According to the different situation of the enterprise, this paper puts forward the following measures to help enterprises improve.

① Increase the enterprise's official microblog exposure, propagandize enterprise microblog through many ways.
According to the "2015 electric business microblog development report" released by sina microblog, the current number of active electric business enterprises in sina microblog account for a total of more than 1600. If you want to highlight your own business account in so many official accounts, you must understand the main way for users to pay attention to enterprise. Seize the basic demands for user to follow the enterprise.

The sina "2015 microblog user development report" shows, users follow the enterprise's microblog primarily through three ways: advertising, others forwarding and corporate propaganda. Besides, Baird [1] thought the establishment of social customer relationship, seized the customer psychology, and the use of interactive communication strategy were also important. In addition to the above measures, the enterprise can also increase the promotion in the major holidays and special time point.

② Pay attention to negative emotions, encourage the positive emotions, and guide the neutral mood.
Chatterjee [7] said that consumers always believe the negative information have more critical value than the positive information when they make purchase decisions or scan the comments. Therefore, they always have greater degree of dependence on the negative information. Through microblog to guide consumers is the main reason for some enterprise to do the microblog marketing. The negative

comments may cause certain extent damage to the product of enterprise. Therefore, some enterprises must pay much attention to the negative emotion and know the origin of it. That can help the enterprise to eliminate the negative emotion.

For the positive emotions, the enterprise should take the initiative to interact with the followers and encourage users to maintain this positive emotion, then to increase the reputation of enterprise. For the neutral emotions, the enterprise should show its advantage to consumers with the objective data.

③ Both active and passive comments should be concerned
Most followers want to get reply or attention when they give comments to the blogger. If bloggers can timely find this kind of behaviors of followers, and then give timely response, that can largely reduce the negative emotions and increase the positive interaction with followers.

④ Diversify the types of microblog
Zhang [19] believed that the types of microblog can to a certain extent influence the browsing and forwarding of fans. Generally the type of blog is divided into: pictures, short chain, video, text, or a mixture of the four types. In order to maintain the freshness of microblog, the enterprise should cross use multiple types of posts to give user a new experience.

4 Conclusions

This paper through the literature research, put forward the evaluation indicators of the microblog marketing effect in enterprises, and set up a set of comprehensive evaluation system.

Combining with the method of data envelopment analysis (DEA), this paper proposes the evaluation model, with input and output of the quantitative evaluation method to evaluate the microblog marketing effect between different enterprises. In the empirical part, this paper selects 50 different enterprises to verify the feasibility of the model. According to the objective data, this paper makes a comparison of different enterprise marketing effect, and provides a feasible improvement measures for it.

In fact, this paper also has certain limitation in the selection of indicators. The article does not use the enterprise actual sales as the evaluation index. Besides, this paper only studies the different enterprises in the short period. The future research can increase the time and carry on deeper research on microblog marketing effects.

Acknowledgements. The authors acknowledge the financial support National Natural Science Foundation Project of China, project (No. 71372189), and Sichuan University Fundamental Research Funds for the Central Universities Project (No. skqy201406). The author is grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

1. Baird CH, Parasnis G (2011) From social media to social customer relationship management. Strategy Leadersh 39(5):30–37
2. Barry AE, Bates AM et al (2016) Alcohol marketing on twitter and instagram: evidence of directly advertising to youth/adolescents. Alcohol Alcohol 51(4):487–492
3. Bi L (2013) Enterprise micro-blogging marketing effect evaluation model and empirical research based on micro-blogging dissemination of information flow. J Intell 7:67–71
4. Bressler BRG (2010) The measurement of productive efficiency. In: Western Farm Economic Association, Proceedings, vol 3, pp 253–290
5. Chang YT, Yu H, Lu HP (2015) Persuasive messages, popularity cohesion, and message diffusion in social media marketing. J Bus Res 68(4):777–782
6. Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444
7. Chatterjee P (2001) Online reviews: do consumers use them? Adv Consum Res 28:133–139
8. Chen H (2015) College-aged young consumers interpretation of twitter and marketing information on twitter. Young Consum 16(2):208–221
9. Clark EM, Jones CA et al (2016) Vaporous marketing: uncovering pervasive electronic cigarette advertisements on twitter. Plos One 11(7):e0157304
10. CNNIC (2016) Statistical report on internet development 38th China internet network
11. Cooper WW, Seiford LM, Tone K (2001) Data envelopment analysis: a comprehensive text with models, applications, references and dea-solver software. Springer, Heidelberg
12. Kafeza E, Makris C, Vikatos P (2016) Marketing campaigns in twitter using a pattern based diffusion policy. In: 2016 IEEE International Congress on Big Data (BigData Congress), pp 125–132
13. Leung XY (2015) The marketing effectiveness of social media in the hotel industry: a comparison of facebook and twitter. J Hospitality Tourism Res 39(2):147–169
14. Liu G, Amini MH et al (2016) Best practices for online marketing in twitter: an experimental study. In: IEEE international conference on electro information technology, pp 504–509
15. Park SB, Ok CM, Chae BK (2016) Using twitter data for cruise tourism marketing and research. Travel Tourism Mark 33(6):885–898
16. Saulles MD (2015) Push and pull approaches to using twitter as a marketing tool. In: Proceedings of the European conference on social media, pp 105–111
17. Sun YL (2008) Evaluation on the capacity of agricultural sustainable development in sichuan province based on dea method. Soft Sci 6:100–103 (in Chinese)
18. Xue JP, Yu WP, Niu YG (2013) Research on brand communication effect of e-commerce enterprises' microblogging - a case study of 51buy. com's microblogging. Soft Sci 12:67–71 (in Chinese)
19. Zhang L, Peng TQ et al (2014) Content or context: which matters more in information processing on microblogging sites. Comput Hum Behav 31(1):242–249
20. Zhu YQ, Chen HG (2015) Social media and human need satisfaction: implications for social media marketing. Bus Horiz 58(3):335–345

Accounting for Clustering and Non-ignorable Missingness in Binomial Responses: Application to the Canadian National Survey on Child Safety Seat

Syed Ejaz Ahmed[1(✉)], Abdulkadir A. Hussein[2], Anne Snowdon[2], and Yiwen Tang[3]

[1] Brock University, 500 Glenridge Avenue, St. Catharines, ON L2S 3A1, Canada
babak_rezaee@aut.ac.ir
[2] University of Windsor, 401 Sunset Avenue, Windsor, ON N9B 3P4, Canada
[3] Cancer Care Ontario, 505 University Avenue, Toronto, ON M5G 1X3, Canada

Abstract. Motivated by a Canadian national survey on child safety seat use, we propose a procedure for incorporating non-ignorable missingness in a binary response using logistic regression model with random effects. The proposed method applies an expectation maximization (EM) algorithm to the Penalized quasi-likelihood of the artificially completed data. We provide closed form formulae for estimating the covariance of the regression coefficients as well as the odds ratios of interest and probabilities of missingness. The proposed algorithm can be easily implemented by using familiar statistical packages that fit generalized linear mixed models. The method is illustrated by applying it to data from Canadian national survey on child safety seat use.

Keywords: Generalized linear mixed model · EM algorithm · Non-ignorable missingness · Data augmentation · Logistic regression · Child safety

1 Introduction

Subjects are likely to decline participation in a survey if the variables measured are socially sensitive. This, raises the issue of non-ignorable missingness whereby, the missignness mechanism depends on the value of the response itself. The current research was motivated by data from a large Canadian survey in which 200 retail parking lots across Canada were random chosen and vehicles entering to the parking lot were asked consent to participate in a survey which measured the correct use of child safety seats in vehicles. Age, height, weight and type of car safety seat used were recorded for up to three children in the participating vehicles. By using these variables, a binary correct use variable was then defined.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_73

Unfortunately, drivers of more than 50% of the vehicles refused to participate in the survey. Since correct uses of children in the same participating vehicle are correlated, in addition to the missing responses, the data presented also a cluster correlation feature.

Ibrahim et al. [3] have developed an EM algorithm for the generalized linear mixed models (GLMM) with normal random effects and non-ignorable missingness in the responses by using MCMC sampling techniques. Ibrahim et al. [3] modeled the respondent data as a function of covariates by using GLMM with logit link and likewise they modeled the non-response data as a function of the same set of covariates as well as the response itself. Thus, they captured the non-ignorable nature of the non-response through a logit link as well. The authors then, proposed an EM (Expectation Maximization) algorithm along with MCMC sampling for estimating the various parameters of the model.

Ibrahim and Molenberghs [8] have recently given an exhaustive review of methods for accommodating non-ignorable missingness in longitudinal data. Also, Huang and Carriere [2] discussed accommodation of incomplete responses data in small sample longitudinal data.

In this paper, we propose a direct method which does not require Monte Carlo sampling. The method proposed here is based on the penalized quasi-likelihood of Breslow and Clayton [1] (PQL) which utilizes Laplacian method of integral approximation of integrals. We use GLMM to model the respondent data, thus accommodating the clustering. On the other hand, we model the dependence of the non-response on covariates and on the true values of the response itself through generalized linear model GLM as was done in Ibrahim et al. [3]. However, we use the PQL method of Breslow and Clayton via Laplace transform and Taylor series approximation [1,6] in order to eliminate the random effects. We then apply an EM algorithm to estimate the regression coefficients and Louis's method [4] to obtain formulae for the their covariances. Following Zeger et al. [9], and by using delta method, we compute marginal probabilities of the response as well as odd ratios and probabilities of non-response and their respective confidence intervals. The proposed method can be easily implemented through familiar statistical software that fit GLMM. In Sect. 2 we give technical details of the model and the algorithm used for fitting it as well as formulae for variances of the various parameters of interest. In Sect. 3 we apply the method to the Canadian national survey on child safety seat use. Finally, in Sect. 4, we give some discussions.

2 The Proposed Method

2.1 Model Details

Let y_i be a binomial variable with size m_i and success probability p_i. Suppose that the complete data consists of a sample $\mathbf{y} = (y_1, \cdots, y_n)'$, from such distribution. Given an n-dimensional vector $b = (b_1, \cdots, b_n)'$ of random effects, we assume that $y_i \mid b$ follows generalized linear mixed model (GLMM), that is

$$y_i \mid b \sim \text{Binomial}\,(m_i, p_i)\,,$$
$$b \sim MVN\,(0, \mathbf{D})\,,$$

where MVN stands for multivariate normal distribution with covariance matrix $\mathbf{D}$, $n \times n$ diagonal matrix with diagonal elements σ^2.

Let $\beta = (\beta_0, \beta_1, \cdots, \beta_p)'$ be a vector of fixed effects regression coefficients associated with the vector of covariates $x_i = (1, x_{i1}, \cdots, x_{ip})'$. Let also, t_i be an $1 \times n$ vector of explanatory variables for the random effects associated with the ith observation. The ith component of t_i is 1 and the rest are all 0. Such vectors can be compacted for all $i = 1, \cdots, n$ in an $n \times n$ identity matrix $T = (t_1, \cdots, t_n)$, which is simply the design matrix associated with the random effects. By standard theory of GLMM, $y_i \mid b$ has the density

$$f(y_i \mid x_i, \beta, \sigma^2, b) = \exp\left\{ y_i \log(p_i/(1-p_i)) + m_i \log(1-p_i) + \log\binom{m_i}{y_i} \right\}, \tag{1}$$

with the first two moments given by

$$E(Y_i \mid b) = \mu_i(b) = m_i p_i \tag{2}$$
$$\text{Var}(Y_i \mid b) = m_i p_i (1-p_i). \tag{3}$$

We use the logit link function to relate covariates (both fixed and random) and the p_i as follows

$$\log(p_i/(1-p_i)) = \eta_i(b) = x_i'\beta + t_i'b = x_i'\beta + b_i. \tag{4}$$

We introduce a missing data indicator variable r_i for each y_i observation in the complete data,

$$r_i = \begin{cases} 1, & \text{if } y_i \text{ is missing;} \\ 0, & \text{if } y_i \text{ is observed.} \end{cases}$$

Now, as in [3], we assume that the missingness indicators, $\mathbf{r} = (r_1, \cdots, r_n)'$, are independent of the random effects, b_i, but depend on the covariates x_i as well as on the value of the response, y_i, through a logistic regression model. Thus, by setting $z_i = (x_i, y_i)'$ to be the vector of explanatory variables for the non-response model with corresponding regression coefficients, $\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_{p+1})'$, we can write the density of the r_i as

$$f(r_i | z_i, \alpha) = \exp\left\{ r_i z_i'\alpha - \log\left(1 + \exp\{z_i'\alpha\}\right)\right\},$$

where the logit link function is again used to connect covariates and probability of non-response $\pi_i = \Pr(r_i = 1 \mid x_i, y_i)$ as follows

$$\log(\pi_i/(1-\pi_i)) = z_i'\alpha.$$

The log-marginal likelihood of $\mathbf{y}$ can now be written as

$$\log f(\mathbf{y}|x, \beta, \sigma^2) = \log \int f(\mathbf{y} \mid x, \beta, \sigma^2, b) f(b) db = \log \int e^{h(b)} db, \tag{5}$$

where $h(b) = \log f(\mathbf{y} \mid x, \beta, \sigma^2, b) + \log f(b)$.

Since the random effects are assumed to be normal, we have

$$\log f(b) = -\frac{1}{2}b'\mathbf{D}^{-1}b - \frac{n}{2}\log 2\pi - \frac{1}{2}\log|\mathbf{D}|$$

and hence,

$$\begin{aligned} h(b) &= \log f(\mathbf{y} \mid x, \beta, \sigma^2, b) + \log f(b) \\ &= \log f(\mathbf{y} \mid x, \beta, \sigma^2, b) - \frac{1}{2}b'\mathbf{D}^{-1}b - \frac{n}{2}\log 2\pi - \frac{1}{2}\log|\mathbf{D}|. \end{aligned}$$

The log-likelihood in Eq. (5) is not easy to deal with as it involves a high dimensional integral. To simplify it, we apply the Laplace approximation of [1,6,7], based on a second-order Taylor series expansion,

$$\log \int e^{h(b)}db \doteq h(\hat{b}) + \frac{n}{2}\log 2\pi - \frac{1}{2}\log\left|-\frac{\partial^2 h(b)}{\partial b \partial b'}\right|, \tag{6}$$

where $\hat{b}$ is solution to

$$\left.\frac{\partial h(b)}{\partial b}\right|_{b=\hat{b}} \left. T'(\mathbf{y} - \mu(b)) - \mathbf{D}^{-1}b\right|_{b=\hat{b}} = 0.$$

It can be shown that [6],

$$-\frac{\partial^2 h(b)}{\partial b \partial b'} = (T'WT\mathbf{D} + I)\mathbf{D}^{-1}, \tag{7}$$

where W is the $n \times n$ diagonal matrix with diagonal element $w_i = m_i p_i(1 - p_i)$.

By using Eqs. (6) and (7), the log-likelihood in Eq. (5) can be approximated by

$$\log f(\mathbf{y}|x, \beta, \sigma^2) \doteq \log f(\mathbf{y} \mid x, \beta, \sigma^2, \hat{b}) - \frac{1}{2}\hat{b}'\mathbf{D}^{-1}\hat{b} - \frac{1}{2}\log|T'WT\mathbf{D} + I|,$$

or equivalently,

$$\begin{aligned} \log f(\mathbf{y}|x, \beta, \sigma^2) \doteq \sum_{i=1}^{n} \Bigg[& y_i\eta_i(\hat{b}) - m_i \log\left(1 + \exp\left\{\eta_i(\hat{b})\right\}\right) + \log\binom{m_i}{y_i} - \frac{1}{2}\frac{\hat{b}_i^2}{\hat{\sigma}^2} \\ & - \frac{1}{2}\log\left(m_i\left(\exp\left\{\eta_i(\hat{b})\right\} / \left(1 + \exp\left\{\eta_i(\hat{b})\right\}\right)^2\right)\hat{\sigma}^2 + 1\right)\Bigg]. \end{aligned} \tag{8}$$

The log-likelihood so obtained is the penalized quasi-log-likelihood and most of the available statistical software can fit such model.

Finally, the joint log-likelihood of the complete data and the missingness indicator can be written as

$$\begin{aligned} l(\alpha, \beta, \sigma^2|\mathbf{r}, \mathbf{y}, x) &= \log f(\mathbf{r}, \mathbf{y}|x, \alpha, \beta, \sigma^2) \\ &= \sum_{i=1}^{n} \left[\log f(r_i \mid z_i, \alpha) + \log f(y_i \mid x_i, \beta, \sigma^2)\right]. \end{aligned}$$

This implies that the i^{th} observation's contribution to the log-likelihood is

$$l_i(\alpha, \beta, \sigma^2 | r_i, y_i, x_i) = \log f(r_i \mid z_i, \alpha) + \log f(y_i \mid x_i, \beta, \sigma^2)$$

and so, the i^{th} observation's contribution to the expected log-likelihood is

$$E[l_i(\alpha, \beta, \sigma^2 | r_i, y_i, x_i)] = \begin{cases} \sum_{y_i=0}^{m_i} w_i(y_i) l_i(\alpha, \beta, \sigma^2 | r_i, y_i, x_i), & \text{if original } y_i \text{ is missing;} \\ l_i(\alpha, \beta, \sigma^2 | r_i, y_i, x_i), & \text{if original } y_i \text{ is observed,} \end{cases}$$

where the weights $w_i(y_i)$ are

$$w_i(y_i) = \frac{f(r_i \mid z_i, \alpha) f(y_i \mid x_i, \beta, \sigma^2)}{\sum_{y_i=0}^{m_i} f(r_i \mid z_i, \alpha) f(y_i \mid x_i, \beta, \sigma^2)}, \tag{9}$$

and $f(y_i \mid x_i, \beta, \sigma^2)$ is the marginal density of y_i obtained from Eq. (8). Note that if the original y_i is observed, then $w_i(y_i) = 1$. Now, the expected log-likelihood for all n observations in the complete binomial sample is

$$L(\alpha, \beta, \sigma^2) = \sum_{i=1}^{n} \sum_{y_i=0}^{m_i} w_i(y_i) \log f(r_i \mid z_i, \alpha) + \sum_{i=1}^{n} \sum_{y_i=0}^{m_i} w_i(y_i) \log f(y_i \mid x_i, \beta, \sigma^2). \tag{10}$$

This last equality indicates that the model for the complete data and that of the missingness indicator can be separated as in Ibrahim et al. [3], and therefore, estimation of the regression parameters can be carried out separately.

2.2 Estimating the Regression Coefficients

The regression coefficients of the model can be estimated by using Eq. (10). First, we create an augmented data set (y_i, r_i, x_i), in which each missing observation y_i is replaced by its $m_i + 1$ possible values, then we apply the following algorithm on the augmented data set.

(1) Set the initial weights to 1 and fit ordinary weighted GLMM to the augmented (complete) data $(\mathbf{y}, \mathbf{x})$ by using any statistical software that fits GLMM via penalized quasi-likelihood methods, and hence, obtain the starting values of β, b, and σ^2.
(2) Set the initial weights to 1 and fit weighted GLM to the missing indicator data $(\mathbf{r}, \mathbf{z})$ and obtain the starting values of α.
(3) By using (9), update the weights $w_i(y_i)$ based on the current values of β, α, b, and σ^2.
(4) Update β, b, and σ^2 by fitting weighted GLMM to the augmented (complete) data $(\mathbf{y}, \mathbf{x})$ with the current values of $w_i(y_i)$.
(5) Update α by fitting weighted GLM to $(\mathbf{r}, \mathbf{z})$ with the current values of $w_i(y_i)$.
(6) Update the weights $w_i(y_i)$ based on the new β, α, b, and σ^2.
(7) Repeat steps 4–6 until β and α converge.

The final estimates of β, α, σ^2 are obtained upon convergence of the above algorithm. Following Louis [4], the estimated observed Fisher information matrix of the parameters $(\beta, \alpha, \sigma^2)$ can be computed as

$$\mathscr{I}(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2) = Q(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2) - H(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2),$$

where

$$\begin{aligned} H(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2) = & \left[\sum_{i=1}^{n} \sum_{y_i=0}^{m_i} \hat{w}_i(y_i) S_i(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 | r_i, y_i, x_i) S_i(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 | r_i, y_i, x_i)' \right] \\ & - \sum_{i=1}^{n} U_i(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 | r_i, y_i, x_i) U_i(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 | r_i, y_i, x_i)', \\ \hat{w}_i(y_i) = & \frac{f(r_i \mid z_i, \hat{\alpha}) f(y_i \mid x_i, \hat{\beta}, \hat{\sigma}^2)}{\sum_{y_i=0}^{m_i} f(r_i \mid z_i, \hat{\alpha}) f(y_i \mid x_i, \hat{\beta}, \hat{\sigma}^2)}, \\ U_i(\hat{\alpha}, \hat{\beta} | r_i, y_i, x_i) = & \sum_{y_i=0}^{m_i} \hat{w}_i(y_i) S_i(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 | r_i, y_i, x_i). \end{aligned}$$

The score vector $S_i(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 | r_i, y_i, x_i)$ is given by

$$\begin{aligned} S_i(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 | r_i, y_i, x_i) = & \left. \frac{\partial l_i(\alpha, \beta, \sigma^2 | r_i, y_i, x_i)}{\partial(\alpha, \beta, \sigma^2)} \right|_{(\alpha, \beta, \sigma^2) = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2)} \\ = & \left((r_i - \hat{\pi}_i), (r_i - \hat{\pi}_i) x_{i1}, \cdots, (r_i - \hat{\pi}_i) z_{ip}, (r_i - \hat{\pi}_i) y_i, d_i, d_i x_{i1}, \right. \\ & \left. \cdots, d_i x_{ip}, \partial l_i / \partial \sigma^2 \right)', \end{aligned}$$

where

$$\begin{aligned} d_i = & \; y_i - \hat{\mu}_i(\hat{b}) - \frac{m_i \hat{\sigma}^2 \hat{p}_i (1 - \hat{p}_i)(1 - 2\hat{p}_i)}{2(m_i \hat{\sigma}^2 \hat{p}_i (1 - \hat{p}_i) + 1)}, \\ \frac{\partial l_i}{\partial \sigma^2} = & \; \frac{\hat{b}_i^2}{2\left(\hat{\sigma}^2\right)^2} - \frac{m_i \hat{p}_i (1 - \hat{p}_i)}{2(m_i \hat{\sigma}^2 \hat{p}_i (1 - \hat{p}_i) + 1)}. \end{aligned}$$

The matrix
$$Q(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2) = \begin{pmatrix} Q_1(\hat{\alpha}) & 0 \\ 0 & Q_2(\hat{\beta}, \hat{\sigma}^2) \end{pmatrix}$$
is such that the (j, k)th element of $Q_1(\hat{\alpha})$ is given by

$$Q_{1,jk}(\hat{\alpha}) = - \left. \frac{\partial^2 L(\alpha, \beta, \sigma^2)}{\partial \alpha_j \partial \alpha_k} \right|_{\alpha = \hat{\alpha}},$$

and the $(j, k)th$ element of $Q_2(\hat{\beta}, \hat{\sigma}^2)$ is

$$Q_{2,jk}((\hat{\beta}, \hat{\sigma}^2) = - \left. \frac{\partial^2 L(\alpha, \beta, \sigma^2)}{\partial \beta_j \partial \beta_k} \right|_{\beta = \hat{\beta}, \sigma^2 = \hat{\sigma}^2},$$

for $j,k = 0,1,\cdots,p$ and $Q_{2,p+1,p+1}(\hat{\beta},\hat{\sigma}^2) = -\frac{\partial^2 L(\alpha,\beta,\sigma^2)}{\partial\sigma^2\partial\sigma^2}\Big|_{\beta=\hat{\beta},\sigma^2=\hat{\sigma}^2}$. These matrices are, respectively, the usual Fisher and penalized Fisher information matrices associated with the GLM and GLMM models and they can be easily extracted from existing software upon convergence of the EM algorithm described above.

2.3 Computing the Marginal Response Probabilities and Probabilities Of Missingness

In binary response studies, investigators are often interested in the marginal probabilities of the response of interest, $\bar{p}_i = P(Y_i = 1)$. This is, in fact, not easy in a random effects model in which a logit link function has been used, due to the non-linearity of the logit function. A way about this difficulty is to use a Taylor approximation of the type proposed by Zeger et al. [9] in which $\bar{p}_i$ are approximately estimated by

$$\hat{\bar{p}}_i = \hat{E}(y_i) \approx \frac{\exp\left\{\hat{k}x_i'\hat{\beta}\right\}}{1+\exp\left\{\hat{k}x_i'\hat{\beta}\right\}},$$

where $\hat{k} = (1+c^2\hat{\sigma}^2)^{-1/2}$ and $c = 16\sqrt{3}/(15\pi)$. By applying delta method to $\hat{\eta}_i = g(\hat{\beta},\hat{\sigma}^2) = \hat{k}x_i'\hat{\beta}$, we obtain estimators of $\hat{\bar{p}}_i$,

$$\hat{V}(\hat{\eta}_i) \approx \hat{G}'\hat{\Sigma}\hat{G},$$

where $\hat{\Sigma}$ is the estimated asymptotic covariance matrix of $\hat{\beta},\hat{\sigma}^2$, that is the lower right block of $\mathscr{I}^{-1}(\hat{\alpha},\hat{\beta},\hat{\sigma}^2)$, and

$$G = \begin{pmatrix} \frac{\partial g(\beta,\sigma^2)}{\partial\beta}\big|_{\hat{\beta},\hat{\sigma}^2} \\ \frac{\partial g(\beta,\sigma^2)}{\partial\sigma^2}\big|_{\hat{\beta},\hat{\sigma}^2} \end{pmatrix}.$$

An approximate 95% confidence interval for η_i is then given by

$$(\hat{\eta}_i^L,\hat{\eta}_i^U) \equiv \left(\hat{\eta}_i - 1.96\sqrt{\mathrm{Var}(\hat{\eta}_i)}, \hat{\eta}_i + 1.96\sqrt{\mathrm{Var}(\hat{\eta}_i)}\right),$$

whereas an approximate 95% confidence interval for $\bar{p}_i = \exp\{\eta_i\}/(1+\exp\{\eta_i\})$ is

$$\left(\frac{\exp\{\hat{\eta}_i^L\}}{1+\exp\{\hat{\eta}_i^L\}}, \frac{\exp\{\hat{\eta}_i^U\}}{1+\exp\{\hat{\eta}_i^U\}}\right).$$

Finally, $\hat{V}(\hat{\bar{p}}_i)$ can be computed by a second delta method as $\hat{V}ar(\hat{\bar{p}}_i) \approx [\hat{\bar{p}}_i(1-\hat{\bar{p}}_i)]^2\hat{V}ar(\hat{\eta}_i)$. The probabilities of non-response, $\hat{\bar{\pi}}_i$, can be easily obtained from the output of the GLM model fit as these probabilities depend only on the regression parameters α_i.

3 Application

The data analyzed in this example are based on a Canadian child safety seats survey described in Snowdon et al. [5]. The main objective of the survey was to measure whether children traveling in vehicles on Canadian roads are correctly restrained in safety devices that are appropriate for their ages, weights and heights. Drivers of vehicles entering in 200 randomly selected retail parking lots across Canada were asked to participate in the survey. For the vehicles whose drivers agreed to participate in the survey, heights, weights, ages and type of safety restraints used were recorded for up to three child occupants in the vehicle. For vehicles whose drivers refused to participate in the survey, the number of children in the vehicle were recorded. In general, there is no consensus on the definition of correct seat for a child. There are several criteria of correct use based on one or combination of age, weight and height. In this article, we employed correct use definition given in Table 1, based only on child's age. We defined the number of children who are correctly restrained in a vehicle as y_i, assumed

Table 1. Definition of correct use of CSS based on the child's age groups

Child safety seats	Age groups
Rear-facing infant seat	(Age $\leqslant$ 1 year)
Forward-facing infant seat	(1 years $<$ Age $\leqslant$ 4 years)
Booster seat	(4 years $<$ Age $\leqslant$ 9 years)
Seat belt	(Age $>$ 9 years)

Table 2. Sample proportions of correct use of CSS and Marginal probabilities of correct use adjusted for missingnes

Category	Proportion	Estimate	*SE*	95% *CI*	
				Lower	Upper
AB	79.4%	65.87%	0.69%	64.50%	67.20%
BC	79.9%	68.01%	0.77%	66.49%	69.49%
MB	74.6%	60.55%	1.39%	57.79%	63.24%
NB	76.0%	68.35%	0.84%	66.68%	69.97%
NFL	72.9%	64.31%	1.12%	62.09%	66.48%
NS	79.3%	72.64%	1.00%	70.64%	74.56%
NT	70.4%	61.47%	2.62%	56.24%	66.46%
ON	84.1%	69.92%	0.54%	68.85%	70.97%
PEI	66.7%	62.20%	1.85%	58.51%	65.76%
QC	71.0%	60.81%	1.07%	58.69%	62.89%
SK	75.7%	65.10%	1.55%	62.01%	68.07%
Nation	78.6%	67.42%	0.34%	66.75%	68.09%

Table 3. Estimated regression parameters β.

Parameter	Estimate	*SE*	t Value	Pr >| t |
Intercept	0.9202	0.0315	29.2553	0.0000
AB	−0.2032	0.0562	−3.6134	0.0003
BC	−0.0977	0.0627	−1.5582	0.1192
MB	−0.4530	0.0938	−4.8296	0.0000
NB	−0.0803	0.0698	−1.1498	0.2503
NFL	−0.2777	0.0841	−3.3010	0.0010
NS	0.1449	0.0889	1.6304	0.1031
NT	−0.4105	0.1756	−2.3378	0.0194
PEI	−0.3768	0.1357	−2.7761	0.0055
QC	−0.4410	0.0774	−5.6954	0.0000
SK	−0.2400	0.1107	−2.1669	0.0303

Table 4. Odds ratios of correct use of CSS and their 95% confidence intervals for all provinces as compared to Ontario

Province	Odds ratio	95% *CI*	
		Lower	Upper
AB	0.8161	0.7124	0.9349
BC	0.9069	0.7806	1.0537
MB	0.6357	0.5062	0.7985
NB	0.9228	0.7852	1.0845
NFL	0.7575	0.6231	0.9209
NS	1.1559	0.9461	1.4123
NT	0.6633	0.4377	1.0052
PEI	0.6860	0.5077	0.9271
QC	0.6434	0.5363	0.7720
SK	0.7866	0.6048	1.0231

to be Binomial(m_i, p_i), where $m_i = 1, 2, 3$. For vehicles whose drivers refused to participate in the survey, y_i is not observed but only m_i is known. Therefore, our objective is to incorporate this missingness into the estimation of the probabilities of correct use, p_i, and probabilities of missingness, π_i, for the various provinces of Canada. The covariates of interest in this study are the dummy codings for 10 Canadian provinces and territories, $x_{1i}, x_{2i}, \cdots, x_{10i}$. We assumed the province of Ontario to be the reference province. We then applied the EM algorithm proposed in this article and summarized results in Tables 2, 3, 4 and 5. In Table 2, we reported the estimated adjusted marginal probabilities of correct use for the various provinces as well as for the entire Canada. It is clear that the adjusted rates exhibit downward bias as compared to the raw proportions.

Table 5. Estimated regression parameters α.

Parameter	Estimate	*SE*	Z Value	Pr > | *Z* |
Intercept	1.2630	0.0722	17.505	0.0000
AB	0.1908	0.1001	1.907	0.0566
BC	−0.0396	0.1061	−0.374	0.7087
MB	0.2648	0.1820	1.456	0.1455
NB	−0.8039	0.1068	−7.524	0.0000
NFL	−0.7253	0.1310	−5.535	0.0000
NS	−1.2439	0.1306	−9.522	0.0000
NT	−0.1025	0.3148	−0.325	0.7448
PEI	−1.4363	0.2204	−6.515	0.0000
QC	−0.4741	0.1268	−3.739	0.0002
SK	−0.3529	0.1830	−1.929	0.0538
y_i	−0.2964	0.0428	−6.927	0.0000

Table 6. Probabilities of non-response, $\hat{\pi}_0$, $\hat{\pi}_1$, $\hat{\pi}_2$, $\hat{\pi}_3$, respectively, when correct use within vehicle is zero, one, two and three

Category	Raw non response rate	$\hat{\pi}_0$	$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{\pi}_3$
AB	75.95%	81.06%	76.09%	70.29%	63.75%
BC	70.74%	77.27%	71.65%	65.26%	58.28%
MB	79.14%	82.17%	77.41%	71.81%	65.44%
NB	55.10%	61.28%	54.06%	46.66%	39.41%
NFL	59.81%	63.13%	56.00%	48.62%	41.30%
NS	45.27%	50.48%	43.11%	36.04%	29.52%
NT	66.67%	76.14%	70.35%	63.82%	56.74%
ON	74.18%	77.95%	72.44%	66.15%	59.24%
PEI	43.62%	45.68%	38.47%	31.73%	25.68%
QC	65.53%	68.76%	62.07%	54.89%	47.49%
SK	69.83%	71.30%	64.88%	57.86%	50.52%
Nation	68.74%	72.73%	67.14%	61.02%	54.54%

In Tables 3 and 4, we reported the estimated regression coefficients, $\hat{\beta}$, and odds ratios based on these parameters. The provinces of Alberta, Manitoba, Newfoundland, Prince Edward Island and Quebec have significantly lower odds of correct use as compared to Ontario. In Table 5, the missing indicator regression parameters, $\hat{\alpha}$, are reported. We clearly see that the coefficient of y_i, α_{11}, is statistically significant, which indicated that the missingness is non-ignorable. In Table 6, we reported the non-response proportions as well as non-response proportions estimated from the EM model. The probabilities of non-response decrease as the number of correctly restrained children in a vehicle increase. These probabilities are denoted by $\hat{\pi}_i$, where $i = 0, 1, 2, 3$ is the number of correctly restrained children in the vehicle.

4 Conclusion

In this manuscript, we proposed a method for handling non-ignorable missing responses in a clustered binary data. We employed GLMM along with penalized quasi-likelihood estimation method for fitting the parameters of the complete data and GLM for fitting those of the missingness mechanism and combined them via the EM algorithm. Procedures for estimating the marginal probabilities of the response are derived via delta method. The proposed methods are then applied to a Canadian national survey for estimating rates of correct use of child safety seats in vehicles. The proposed methods are attractive in the sense that they only require existing statistical software that have capabilities for fitting GLMM and GLM.

References

1. Breslow NE, Clayton DG (1993) Approximate inference in generalized linear mixed models. J Am Stat Assoc 88(421):9–25
2. Huang R, Carriere KC (2006) Comparison of methods for incomplete repeated measures data analysis in small samples. J Stat Plann Infer 136(1):235–247
3. Liechty JC, Roberts GO (2001) Missing responses in generalised linear mixed models when the missing data mechanism is nonignorable. Biometrika 88(2):551–564
4. Louis TA (1982) Finding the observed information matrix when using the EM algorithm. J Roy Stat Soc 44(2):226–233
5. Snowdon AW, Hussein A et al (2009) Are we there yet? Canada's progress towards achieving road safety vision 2010 for children travelling in vehicles. Int J Inj Control Saf Promot 16(4):231–237
6. Song XK (2007) Correlated data analysis: modeling, analytics, and applications. Springer, New York
7. Tierney L, Kadane JB (1986) Accurate approximations for posterior moments and marginal densities. J Am Stat Assoc 81(393):82–86
8. Ugarte MD (2009) Missing data methods in longitudinal studies: a review. Test 18(1):1–43
9. Zeger SL, Albert PS (1988) Models for longitudinal data: a generalized estimating equation approach. Biometrics 44(4):1049–60

Equity and Sustainability Based Model for Water Resources Planning

Yan Tu[1], Jun Gang[2], Benjamin Lev[3], and Xiaoyang Zhou[4](✉)

[1] School of Management,
Wuhan University of Technology, Wuhan 430070, People's Republic of China
tuyan1988@whut.edu.cn

[2] Sichuan Institute of Building Research,
Chengdu 610081, People's Republic of China

[3] Decision Sciences Department, LeBow College of Business,
Drexel University, Philadelphia, PA 19104, USA

[4] International Business School, Shaanxi Normal University,
Xi'an 710062, People's Republic of China
283509632@qq.com

Abstract. This paper presents a decentralized multi-objective bi-level model for water resources planning, in the upper level of the model, the regional authority as the leader determines to allocate how much water to each sub-area to maximize equity and minimize the environmental pollution, and then the sub-areas in the lower level make decisions on allocating water to water sectors based on different allocation principles with the analysis of relationship between supply and demand to maximize the benefits. In the decision making process, fuzzy random environment is considered to describe the internal uncertainty, an expected value operator is used to deal with the fuzzy random variables. Finally, a case study is used to verify the applicability of the proposed model.

Keywords: Water resources allocation · Equity · Sustanability · Bi-level programming · Fuzzy random variable

1 Introduction

In the water resources allocation system, there exists the interaction between regional authority and sub-areas in a river basin [11,12]. Efficiency and sustainability of water resources allocation often conflict with each other, for without considering ecological water to destroy the sustainability of water resources allocation, other water sectors that can generate economic benefits can obtain more water, the efficiency of water resources allocation will improve [4,5,8]. In addition, to guarantee the efficiency and sustainability of water using will lead the contradiction between two-level decision makers (i.e., regional authority and sub-areas).

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_74

Rogers and Louis [9] pointed out that there were a variety of activities and objectives in a water system as well as complicated supply and demand contradictions, which brought pressures to regional authority and sub-areas managers, because these contradictions and ineffectiveness management limit the economic development and environment protection. Therefore, an equitable and efficient water resources allocation is an important measure to deal with water crisis and improve water management, especially when it is very lack of water.

Based on the above, this paper considers to establish a user-friendly water resources allocation model, with two-level structure, in which the regional authority and sub-areas are the upper level and lower level decision makers respectively, as there are multiple decision makers in the lower level, it is a decentralized bi-level model. In the model, the regional authority as the leader aims to maximize the equity (i.e., minimize the Gini coefficient), the sub-areas as the followers expect to maximize the economic benefits. In addition, practical water resources allocation process faces complex uncertain environments, technology progress and equipment update and climatic change will bring uncertainty to parameters in the water transport process, and it is hard to describe the uncertainty using simple random variable or fuzzy variable, therefore, the fuzzy random variable is considered. Based on the discussion above, a decentralized multi-objective bi-level model with fuzzy random coefficients will be established in this paper.

The reminder of this paper is structured as follows. In Sect. 2, the problem statement is given, including the bi-level structure of the model and the motivation to consider fuzzy random variable. Then an expected decentralized bi-level model is established in Sect. 3. In Sect. 4, a case study is conducted to verify the applicability of the model. Finally a conclusion remark is given in Sect. 5.

2 Problem Statement

For a regional water resources allocation system, in the case of insufficient water supply, the limited water should be allocated optimally to each sub-area, and then be allocated by each sub-area to different water sectors (i.e., ecological water, municipal water, industrial water, agricultural water) [10]. The fair allocation of water can balance the contradiction between water supply and demand, therefore, the regional authority the leader in the upper level expects to maximize the equity of water allocation, in the lower level, the sub-areas hope to maximize the benefits through water allocation based on different allocation principles (i.e., efficiency principle, stress principle, priority principle) with the analysis of relationship between supply and demand. The bi-level structure of the water allocation model is shown in Fig. 1.

The need to address uncertainty in water allocation system is widely recognized, there is a strong motivation for considering the fuzziness and randomness in water resources planning problems [1,3,6]. For example, the loss ratio of water transfer is not fixed because of the effect of many uncertain elements such as lack of historical data, technology progress and equipment update, dynamics of

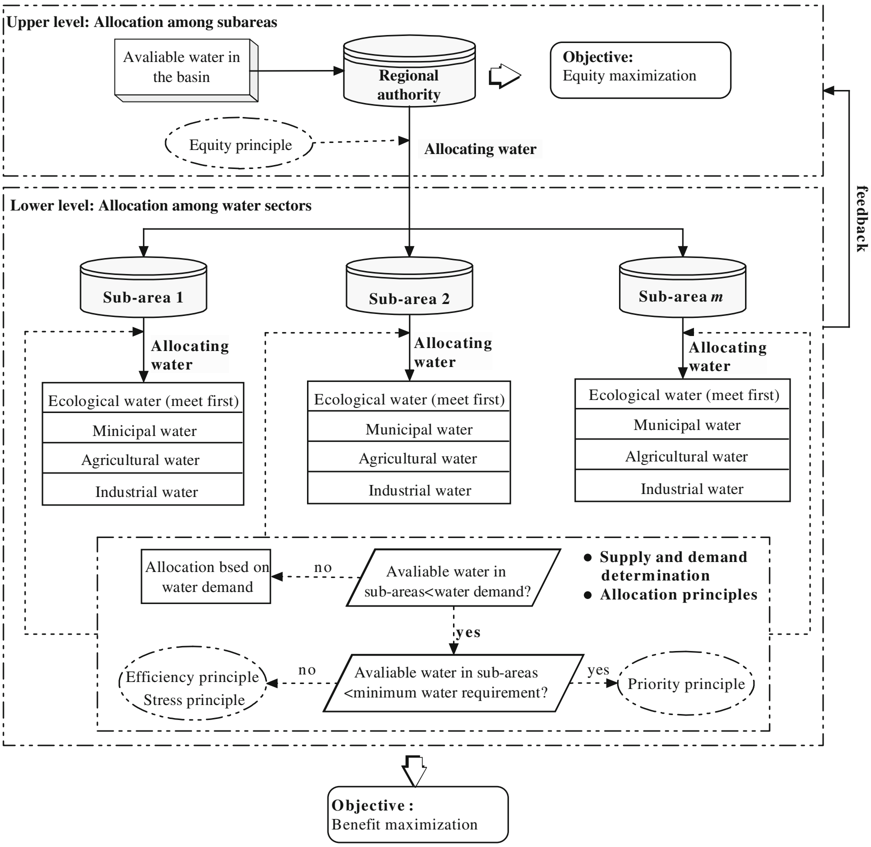


Fig. 1. Bi-level model structure of water allocation

water resources systems, in this case, it is appropriate to use a fuzzy variable to describe the loss ratio of water transfer, which is the fuzziness of the loss ratio of water transfer. The decision making is happen in the future, some experts will be asked to describe the variable using different ranges, which is the randomness of the loss ratio of water transfer. In this study, triangle fuzzy number is used to describe the fuzziness, normal distribution random variable is used to describe the most possible value in the triangle fuzzy number. Therefore, fuzzy random variable is used to describe the two-hold uncertainty of the loss ratio of water transfer, the flowchart of the loss ratio of water transfer as a fuzzy random variable is shown in Fig. 2.

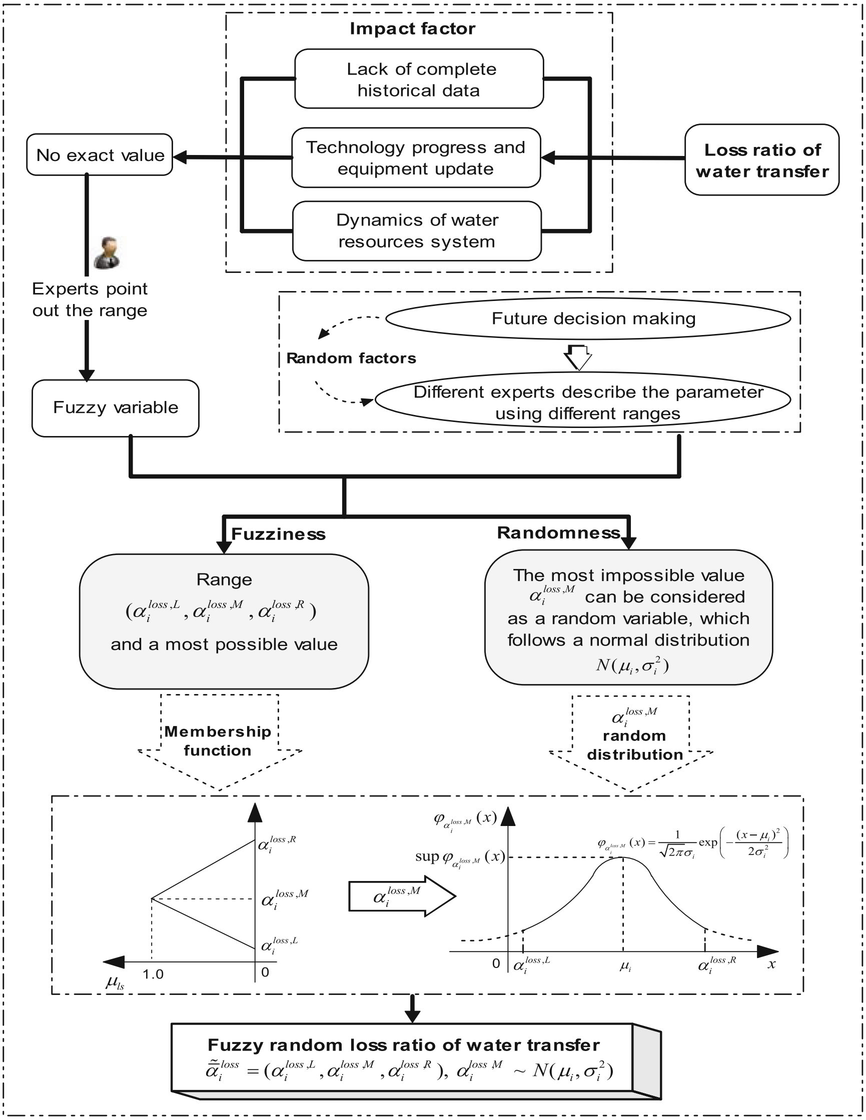


Fig. 2. The flowchart of the loss ratio of water transfer as a fuzzy random variable

3 Modelling

3.1 Assumptions

The following assumptions are considered for constructing the model.

(1) All available water is from a river basin.

(2) The leader (i.e., regional authority) and the followers (i.e., sub-areas) will make rational decision makings for avoiding unsatisfying solutions.
(3) Loss ratio of water transfer is considered as a fuzzy random variable, in which the parameters are determined using data analysis based on historical data and experience.

3.2 Notations

Indices
i: index of sub-area;
j: index of water sector;
k: index of special water sector that requirements should be met;

Parameters
Q: available water of basin;
$e_{i,j}$: economic benefit for using water of water sector in sub-area;
e_i: economic benefit for using water in sub-area;
$o_{i,j}$: amount of the pollutant BOD in unit wastewater discharge of water sector in sub-area;
$w_{i,j}$: wastewater discharge coefficient of water sector in sub-area;
v_i: economic benefit of sub-area;
q_i: water source of sub-area;
$Z_i^{\min}$: minimum capacity of the physical connection between water source and sub-area;
$Z_i^{\max}$: maximum capacity of the physical connection between water source and sub-area;
$S_i^{\max}$: maximum capacity of sub-area;
$d_{i,j}^{\min}$: minimum water requirement of water sector in sub-area;
$d_{i,j}^{\max}$: maximum water requirement of water sector in sub-area;
$ed_{i,j}^{\max}$: minimum ecological water requirement of sub-area;

Fuzzy random variables
$\tilde{\bar{\alpha}}_i^{\text{loss}}$: loss ratio of water transfer from water source to sub-area;

Decision variables
x_i: water allocated to sub-area;
$y_{i,j}$: water allocated to water sector in sub-area

3.3 Model Formulation

Based on the problem statement and the above preparatory work, the decentralized bi-level model for water resources allocation with analysis of supply and demand is established as follows.

$$\max_{\mathbf{x}} G = \frac{1}{2m\sum_{k=1}^{m}\frac{x_k}{e_k}}\sum_{k=1}^{m}\sum_{l=1}^{m}\left|\frac{x_k}{e_k}-\frac{x_l}{e_l}\right| \tag{1}$$

$$\min_{\mathbf{x}} P = \sum_{i=1}^{m}\sum_{j=1}^{n} 0.01 o_{i,j} w_{i,j} y_{i,j} \tag{2}$$

$$\text{s.t. } \sum_{i=1}^{m} x_i \le Q \tag{3}$$

$$(1 - E[\tilde{\bar{\alpha}}_i^{\text{loss}}])x_i + q_i \ge \sum_{k=1}^{t} d_{i,k}^{\min}, \quad \forall i \tag{4}$$

$$z_i^{\min} \le x_i \le z_i^{\max}, \quad \forall i \tag{5}$$

$$(1 - E[\tilde{\bar{\alpha}}_i^{\text{loss}}])x_i + q_i \le s_i^{\max}, \quad \forall i \tag{6}$$

For given $\mathbf{x}, \mathbf{y}$ solves

$$\max_{\mathbf{y}} v_i = \sum_{j=1}^{n} e_{i,j} y_{i,j}, \quad \forall i \tag{7}$$

$$\text{s.t. } d_{i,j}^{\min} \le y_{i,j} \le d_{i,j}^{\max}, \quad \forall i, j \tag{8}$$

$$(1 - E[\tilde{\bar{\alpha}}_i^{\text{loss}}])x_i + q_i - \sum_{j=1}^{n} y_{i,j} \ge ed_i^{\min}, \forall i \tag{9}$$

Equation (1) represents the first objective of the leader (i.e., regional authority), which aims to maximize the equity of the water allocation based on Gini coefficient [13]. Equation (2) represents the second objective of the leader (i.e., regional authority), which aims to minimize the environmental pollution to achieve the sustainability in water using at most. Constraint in Eq. (3) requires the water allocated to all sub-areas should be less than the available water in the river basin. Constraints in Eq. (4) require the water allocated to sub-area should meet the water using of special waters in sub-area. $E[\bullet]$ is the expected value operator proposed in Liu [7]. Constraints in Eq. (5) describe the range of water allocated to sub-area. Constraints in Eq. (6) restrict all water sub-area can get should be less than its maximum capacity. Equations in Eq. (7) represent the objectives of the followers (i.e., sub-areas), which aim to maximize the economic benefit of all the sub-areas. Constraints in Eq. (8) define the range of water allocated to water sector of sub-area. Constraints in Eq. (9) require that the ecological water should be met in each sub-area.

4 Model Analysis

From the decentralized bi-level model, the following results can be obtained.

(1) When $(1 - E[\tilde{\bar{\alpha}}_i^{\text{loss}}])x_i + q_i - ed_i^{\min} < \sum_{j=1}^{3} d_{i,j}^{\min}$, allocation principle: equity principle in the upper level, priority principle in the lower level.

(2) When $(1 - E[\tilde{\bar{\alpha}}_i^{\text{loss}}])x_i + q_i - ed_i^{\min} \geq \sum_{j=1}^{3} d_{i,j}^{\min}$, $x_i \geq \left(\sum_{j=1}^{3} d_{i,j}^{\min} + ed_i^{\min} - q_i\right) / (1 - E[\tilde{\bar{\alpha}}_i^{\text{loss}}])$, and $e_{i1} = e_{i2} = e_{i3}$, allocation principle: equity principle in the upper level, stress principle in the lower level.

(3) When $(1 - E[\tilde{\bar{\alpha}}_i^{\text{loss}}])x_i + q_i - ed_i^{\min} \geq \sum_{j=1}^{3} d_{i,j}^{\min}$, $x_i \geq \left(\sum_{j=1}^{3} d_{i,j}^{\min} + ed_i^{\min} - q_i\right) / (1 - E[\tilde{\bar{\alpha}}_i^{\text{loss}}])$, and $e_{i1} \neq e_{i2} \neq e_{i3}$, allocation principle: equity principle in the upper level, benefit-based principle in the lower level.

5 Case Study

Qujiang basin is taken as a case example to illustrate the proposed model. Qujiang river is the largest tributary of Jialing river basin in the upper reaches of Yangtze, it is across three provinces in southwest China, in this paper, the part of Qujiang basin in Sichuan Province is studied (see Fig. 3). From Fig. 3, the available water is supplied to five sub-areas: Bazhong (BZ), Nanchong (NC), Guang'an (GA), Guangyuan (GY), Dazhou (DZ). The data for the related parameters can be seen in Tables 1, 2 and 3.

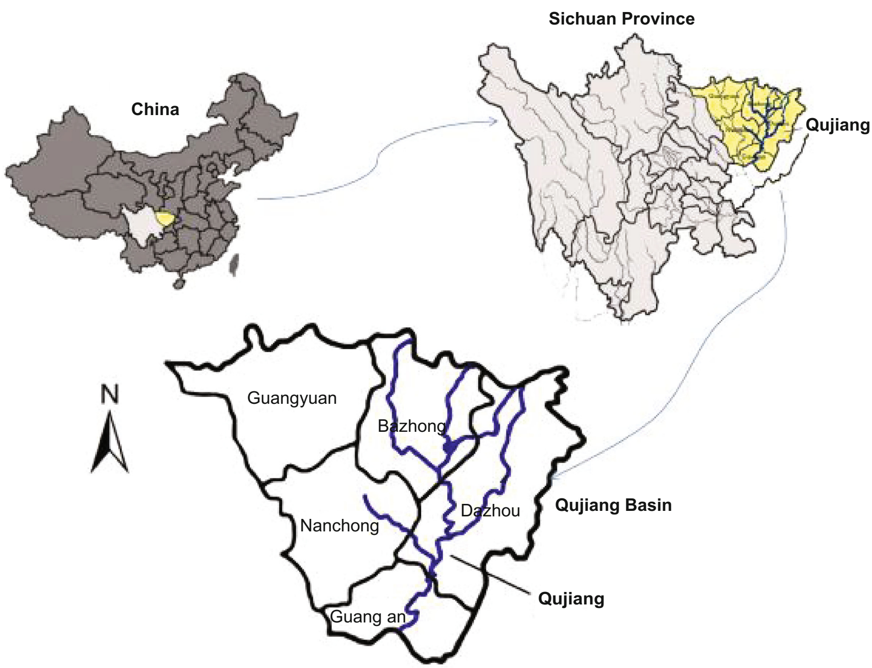


Fig. 3. The map of studied Qujiang river basin

Table 1. Related parameters of 3 water sectors in 5 sub-areas

Sub-area	$\tilde{\bar{\alpha}}_i^{\text{loss}}$	q_i (10^4 m^3)	$s_i^{\max}$ (10^4 m^3)	$z_i^{\max}$ (10^4 m^3)	$z_i^{\min}$ (10^4 m^3)	$e_{i,j}$ (10^4m^3)		
						IND	AGR	MUN
BZ	$(0.42, \alpha_1^{\text{loss}}, 0.47)$ $\alpha_1^{\text{loss}} \sim N(0.45, 0.01)$	42764	50000	53157	47905	60.61	42.55	43.86
NZ	$(0.36, \alpha_2^{\text{loss}}, 0.40)$ $\alpha_2^{\text{loss}} \sim N(0.38, 0.04)$	23870	40000	45320	39514	54.35	34.48	32.26
GA	$(0.28, \alpha_3^{\text{loss}}, 0.32)$ $\alpha_3^{\text{loss}} \sim N(0.30, 0.01)$	33997	40000	46889	18622	67.57	47.62	47.17
GY	$(0.45, \alpha_4^{\text{loss}}, 0.55)$ $\alpha_4^{\text{loss}} \sim N(0.50, 0.09)$	2874	30000	38543	12448	74.07	31.25	37.88
DZ	$(0.29, \alpha_5^{\text{loss}}, 0.35)$ $\alpha_5^{\text{loss}} \sim N(0.32, 0.01)$	115382	60000	67832	26308	86.96	43.11	45.45

Note: IND—Industrial water; AGR—Agricultural water; MUN—Municipal water.

Table 2. Water demand of 3 water sectors in 5 sub-areas (10^4 m^3)

Sub-area	Industrial water		Agricultural water		Municipal water		Ecological water
	$d_{i,1}^{\max}$	$d_{i,1}^{\min}$	$d_{i,2}^{\max}$	$d_{i,2}^{\min}$	$d_{i,3}^{\max}$	$d_{i,3}^{\min}$	
BZ	48848	11715	76465	51917	15510	11410	285
NC	47202	5643	42032	36349	7845	5675	163
GA	45149	16356	44277	33187	9616	6697	149
GY	35988	653	4224	3422	812	659	17
DZ	78174	52444	141104	104172	28092	21702	598

Before solving the proposed model, the expected values of loss ratios of water transfer should be given. For the loss ratios of water transfer are triangle fuzzy random variables, in which the random parameters follow normal distributions. The expected value operator for triangle fuzzy random variables proposed in Xu et al. [12] can be applied to deal with fuzzy random loss ratios of water transfer. Thus, the model can be subsequently processed with the expected value of loss ratios of water transfer.

For dealing with the multiple objectives in the upper level, the weighted sum method based on satisfactory degree proposed by Gang et al. [2] is applied, the

Table 3. Water using pollution parameters of 3 water sectors in 5 sub-areas

Sub-area	Industrial water		Agricultural water		Municipal water	
	$o_{i,1}$	$w_{i,1}$	$o_{i,2}$	$w_{i,2}$	$o_{i,3}$	$w_{i,3}$
BZ	250	0.75	650	0.45	120	0.9
NC	300	0.85	850	0.55	160	0.8
GA	275	0.8	750	0.5	140	0.95
GY	325	0.7	1000	0.45	150	0.85
DZ	200	0.9	800	0.6	170	0.75

Table 4. Water allocation results (10^4 m^3)

Sub-area	Industrial water	Agricultural water	Municipal water	Water allocated by authority
BZ	5550	51917	11410	47955
NC	6050	36349	5675	39364
GA	21858	33187	6697	39792
GY	4850	3322	659	12298
DZ	23289	104172	21702	50490

weights are set as: for the first objective (i.e., the equity objective), for the second objective (i.e., the pollution control objective). Based on the data in Tables 1, 2 and 3, the water allocation planning of this case can be shown in Table 4.

From the results in Table 4, we know that the total available water in Qujiang basin (i.e., 4089×10^6 m^3) is less than the total minimum demand in all sub-areas, but is more than the total minimum demand for agricultural sectors and municipal sectors, which can be known that the water shortage situation is serious currently in the basin. In this paper, the regional authority makes decisions on allocating water based on equity (mainly) and pollution control. Through the comparison of the results with the practical allocation, it can be found that there is scarce water, the minimum demand for agricultural and municipal water sectors of all sub-areas can be satisfied using the proposed model. However, for more economic benefits, more water is allocated for industrial sectors, and the minimum demand for agricultural and minimal using cannot be satisfied in practice, which is not reasonable for the steady and development of the river basin. What's more, the decision making of the proposed model can effectively control the pollution under the premise of guaranteeing equity.

6 Conclusion

This paper proposed a decentralized multi-objective bi-level model considering equity, economic benefit and ecological environment for water resources allocation. In the upper level of the model, the regional authority aimed to achieve two objectives of equity and environmental pollution control by allocating water to sub-areas, in the lower level of the model, the sub-areas expected to get the maximal economic benefit by allocating water to different water sectors. In the model, fuzzy random loss ratio of water transfer was considered. Finally, a case study was conducted to verify the applicability of the proposed model. The comprehensive research on equity, sustainability and effectiveness of water resources allocation under different uncertain environments will be the future research direction.

Acknowledgement. This research was supported by the National Natural Science Foundation of China (Grant No. 71401093), the Fundamental Research Funds for the Central Universities (WUT: 2016VI002), the Fundamental Research Funds for the Central Universities (Grant No. 14SZYB08), Soft Science Research Project of Shaanxi Province (Grant No. 2016KRM089) and Research Center for Systems Science & Enterprise Development (Grant No. Xq16B01).

References

1. Ajami NK, Hornberger GM, Sunding DL (2008) Sustainable water resource management under hydrological uncertainty. Water Resourc Res 44(11):2276–2283
2. Gang J, Tu Y et al (2014) A multi-objective bi-level location planning problem for stone industrial parks. Comput Oper Res 56:8–21
3. Guo P, Huang GH et al (2010) A two-stage programming approach for water resources management under randomness and fuzziness. Environ Model Softw 25:1573–1581
4. Hu Z, Chen Y et al (2016) Optimal allocation of regional water resources: from a perspective of equity–efficiency tradeoff. Resour Conserv Recycl 109:102–113
5. Lartigue C (2015) Efficient use of water resources for sustainability. Springer, Switzerland
6. Li YP, Liu J, Huang GH (2014) A hybrid fuzzy-stochastic programming method for water trading within an agricultural system. Agric Syst 123(2):71–83
7. Liu B (2004) Uncertainty theory: an introduction to its axiomatic foundations. Springer, Berlin
8. Ni J, Xu J, Zhang M (2016) Constructed wetland planning-based bi-level optimization to balance the watershed ecosystem and economic development: a case study at the Chaohu lake watershed, China. Ecol Eng 97:106–121
9. Rogers JW, Louis GE (2007) A financial resource allocation model for regional water systems. Int Trans Oper Res 14(1):25–37
10. Song WZ, Yuan Y et al (2016) Rule-based water resource allocation in the Central Guizhou Province, China. Ecol Eng 87:194–202
11. Tu Y, Zhou X et al (2015) Administrative and market-based allocation mechanism for regional water resources planning. Resour Conserv Recycl 95:156–173
12. Xu J, Tu Y, Zeng Z (2012) Bilevel optimization of regional water resources allocation problem under fuzzy random environment. J Water Resour Planning Manage 139(3):246–264
13. Yitzhaki S (1979) Relative deprivation and the Gini coefficient: reply. Q J Econ 93(2):321–324

SCADA and Artificial Neural Networks for Maintenance Management

Alberto Pliego Marugán(✉) and Fausto Pedro García Márquez

Ingenium Research Group, Universidad Castilla-La Mancha, 13071 Ciudad Real, Spain
alberto.pliego@uclm.es

Abstract. Nowadays, the reliability of the wind turbines is essential to ensure the efficiency and the benefits of the wind energy. The SCADA system installed in a wind turbine generates lot of data that need to be processed. The information obtained from these data can be used for improving the operation and management, obtaining more reliable systems. The SCADA systems operate through different control rules that are predefined. However, a static control of the wind turbine can generate a miscorrelation between the control and the real conditions of the wind turbine. For example, two wind turbines can be separated several kilometers in the same wind farm, therefore, the operation conditions must be different and the control strategy should not be unique. This research work presents a method based on neural networks for a dynamic generation of the control strategy. The method suggests that the thresholds used for generating alarms can vary and, therefore, the control of the wind turbine will be adapted to each specific wind turbine.

Keywords: Wind turbine · Reliability · SCADA systems · Advanced control analytics · Artificial neural networks

1 Introduction

The wind energy is currently the most important renewable energy, the capacity installed currently is more than 420 GW and it is estimated to be more than 1000GW in 2030 [3]. The maintenance and operation costs of conventional wind turbines are 12% of the total costs, but the wind energy is evolving towards the offshore location. It causes an important increase of these costs, being for offshore wind farms around 23% of the total costs [12].

SCADA systems are widely introduced in wind turbines (WTs) due to their effectiveness has been proved in other industries for detection and diagnostics of failures [9,11,16]. They are presented as an inexpensive and optimal solution [20] to control feedback for the health monitoring while reducing the operation and management costs [19]. Nevertheless, they also present some minor disadvantages due to the operational or reliability conditions [14,21]. These systems consider a large amount of measurements such as temperatures or wind and energy conversion parameters [18]. Data have raised considerable interest in different areas, e.g. wind power forecasting [17], production assessment [22] or fault detection [4,6,8,10].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_75

In the case of WTs, the introduction of SCADA systems verifies the efficiency when their components are deteriorated. This degradation can indicate problems of different nature such as misalignments in the drive-train, friction caused by bearing or gear faults. The basic elements of the performance monitoring consist on a first collection of raw values by the sensors. After the application of the appropriate filters, anomalies are detected. Finally, a diagnosis will be provided. The anomaly detection includes a series of techniques that range from simple threshold checks to statistical analyses [5,7,15].

It has been demonstrated that the WT suffers a gradual loss of production. Figure 1 shows that the power has a decrease year by year (this study has been developed by Sheffield university in the OPTIMUS project [2]). The SCADA data has been considered over 5 years and the wind speed-power curve has been estimated.

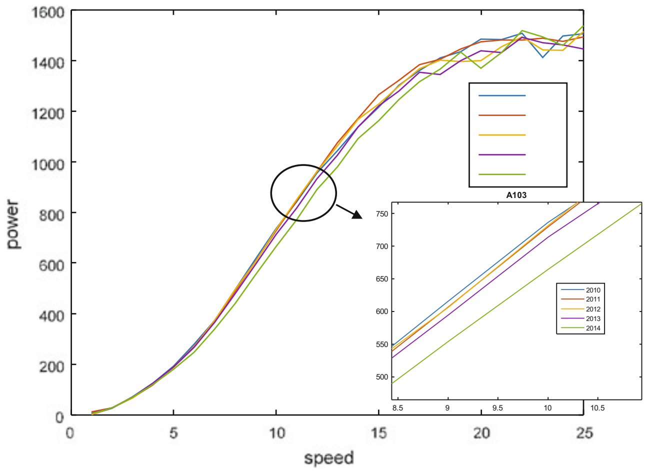


Fig. 1. Loss of power of a WT over the time

This paper present a signal processing considers the minor changes in the behavior of the WT. Some alarms will be affected by the decrease of the power shown in Fig. 1. The following alarms will be activated considering the power as a cause.

- Activation of the ice safe mode: One of the statements considered for this alarm is that the power is low for the measured wind.
- High aerodynamic deterioration: It is considered for low power with high wind.

If the power presented a minor reduction, then some false alarms would be activated because the SCADA was prepared for a higher power. However, if the control laws are dynamic, this problem will be solved.

2 Proposed Method

The method proposed aims to improve the control of the system by following a supervised iterative process. The objective is to determine if the SCADA control is coherent

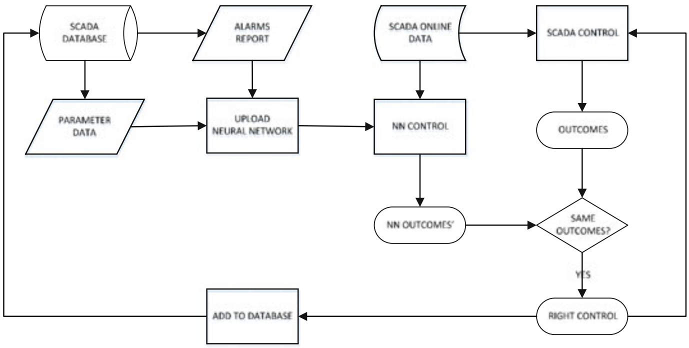


Fig. 2. Flowchart of the method proposed

with the historical data and, therefore, if the system has changed. The main capability of the method is to find differences in the behavior of the system over the time. Figure 2 shows a flowchart of the method proposed.

The iterative process starts with a database of the historical SCADA data. In this paper, two different types of data are considered: the value of the measured parameters and an alarm report. This data is employed to build an artificial neural network (NN) with a supervised training. The values of the parameters are employed as the inputs and the alarms will represent the outputs of the NN.

NN are used in problems that cannot be formulated as an exact method or an analytical solution. NN learns by itself and provides a good solution for the problem simulating the biological neurons in a reasonable time. An artificial NN consists of neurons that are simple processing units and weighted connections between those neurons [1].

It can be used for improving the control of the WT when the NN that defines the logic of the system has been generated. The SCADA system will provide online data of the condition of the WT. These data will be processed following two different evaluations: the SCADA system processing and the NN processing. Both outcomes will be compared. If the outcomes are equal, then the data will be added to the database, there for the NN processing will became an adaptive process.

3 Case Study

The European Optimus project [2] has provided the SCADA data used in this work. It measures a lot of parameters, but after a filtering process, only 34 parameters (see Table 1) have been considered for the analysis presented hereby. The values of this parameters will be considered as inputs for the NN.

Basically, the NN receives a dataset and make a training process to recognize several patterns. The training process fits the different weights to provide the output. If the output is known, then the training is defined as supervised, otherwise, it is called unsupervised training. The condition of the WT corresponds to the desired outputs of the NN. The data used to design the NN is divided in following groups:

Table 1. Inputs and outputs for neural network

No	Signal	No	Signal
1	General accumulator blade 1 pressure	18	Environmental temperature
2	General accumulator blade 2 pressure	19	Drive end side generator bearing temperature
3	General accumulator blade 3 pressure	20	Non-drive end side generator bearing temperature
4	Phi cosine	21	Generator winding temperature
5	Turbulence level	22	Nacelle temperature
6	Oscillation level	23	Lower gearbox radiator
7	Vibration level	24	Upper gearbox radiator
8	Pitch 1 angle	25	Gearbox bearing temperature
9	Pitch 2 angle	26	Transformer 1 temperature
10	Pitch 3 angle	27	Transformer 2 temperature
11	Active power	28	Transformer 3 temperature
12	General accumulator pressure	29	Grid voltage
13	Brake pressure	30	Total reactive power
14	Hydraulic group pressure	31	Generator speed
15	SP pitch angle	32	Rotor speed
16	Hydraulic group oil temperature	33	Wind speed
17	Gearbox oil temperature	34	Yaw

- Training set: Around 75% of the total amount of data.
- Validation set: Around 15% of the total amount of data.
- Testing set: Around 15% of the total amount of data.

The NN is expected to be able to learn from data and predict the output when a generic input is considered.

The output of the NN will be specified by a Report of Alarms generated by the SCADA [2]. The data available consists of a serial of alarms registered from 01/10/2012 to 15/05/2015, where there are different types of alarms identified by a specific code. The codes are not indicated for reasons of confidentiality. In this work, only 5 different scenarios have been considered to simplify the example. These scenarios correspond to 4 different alarms and the absence of alarms.

Therefore, the NN will be created following the structure shown in Table 2.

The design of the NN has been carried out using the geometrical pyramid rule [13]. Figure 3 shows the structure of the NN built. This rule express an approximation of the number of neurons h at the hidden layer. It is defined by:

$$h = \sqrt{m \times n},$$

where m is the number of elements of each input and n is the number of possible outputs, being $m = 34$, $n = 5$, and therefore, $h = 13.03$.

Figure 4 shows the statistics of the NN generated by using the SCADA data. This is a global confusion matrix and corresponds to the sum of the training, validation and testing sets. The rows of the matrix show the outcomes of the NN for the dataset used.

Table 2. Inputs and outputs for NN

	Input data			Output data				
	1	⋯	34	Alarm 1	Alarm 2	Alarm 3	Alarm 4	None Alarm
Date 1	a_1^1	$a_1^{\cdots}$	a_{1j}^{34}	1	0	0	0	0
Date 2	a_2^1	$a_2^{\cdots}$	a_{2j}^{34}	0	1	0	0	0
Date 3	a_3^1	$a_3^{\cdots}$	a_{3j}^{34}	0	0	1	0	0
Date 4	a_4^1	$a_4^{\cdots}$	a_{4j}^{34}	0	0	0	1	0
Date 5	a_5^1	$a_5^{\cdots}$	a_5^{34}	0	0	0	0	0

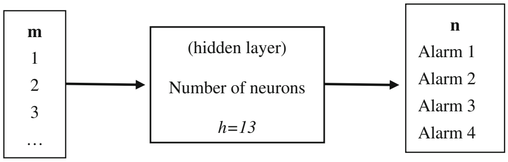


Fig. 3. Structure of the neural network

Output \ Target	1	2	3	4	5	6
1	28 28.6%	0 0.0%	0 0.0%	0 0.0%	1 1.0%	96.6% 3.4%
2	0 0.0%	9 9.2%	1 1.0%	0 0.0%	1 1.0%	81.8% 18.2%
3	0 0.0%	0 0.0%	9 9.2%	0 0.0%	1 1.0%	90% 10%
4	0 0.0%	0 0.0%	0 0.0%	9 9.2%	0 0.0%	100% 0.0%
5	0 0.0%	2 2.0%	1 1.0%	0 0.0%	36 36.7%	92.3% 7.7%
6	100% 0.0%	81.8% 18.2%	81.8% 18.2%	100% 0.0%	92.3% 7.7%	92.9 7.1%

Fig. 4. Results of the neural network: confusion matrix

The columns show the real output, i.e. the output established by the alarms. The diagonal of this matrix (grey cells) contains the desired solutions. The sixth row and the sixth column show a summary of the rights (green percentage) and wrongs (red percentage).

The outcomes of the NN agree more than 90% with the alarms that the SCADA system generates. The method proposed establishes that the control of the SCADA needs to be uploaded when the alarms cannot be predicted by the NN with enough accuracy. If the SCADA control and this NN have the same response, then the data processed will be added to the database. Therefore, a "healthy" dataset will be created and the control will be adapted to the real conditions over the time.

4 Alarm Prediction

The method proposed here by also allows for predicting a possible alarm. In the previous Sect. 3, a NN was generated by inputting the data that the SCADA system employs for generating the alarms. In this section, the input will be the data obtained before the alarm is activated, i.e. the objective is to predict an alarm before the SCADA system generates it. The NN is performed using the same techniques that in previous section. In this case the NN will be designed to distinguish whether alarm will be activated or not. Figure 5 shows the scheme used for the prediction. The inputs employed for performing the NN are the dataset collected before the alarms are activated.

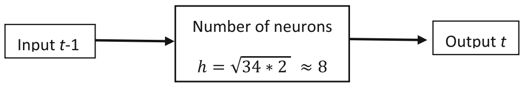


Fig. 5. Structure of the second neural network

Following the structure of Fig. 5, a NN has been performed obtaining the results shown in Fig. 6. This confusion matrix shows that the NN can predict a 20.1% of the alarms. Despite it is a low percentage when the NN suggests an alarm, it has a success of 62.1%.

Output \ Target	1	2	
1	5580 96.5%	143 2.5%	97.5% 2.5%
2	22 0.4%	36 0.6%	62.1% 37.9%
	99.6% 0.4%	20.1% 79.9%	97.1% 2.9%

Fig. 6. Results of the second neural network: confusion matrix

This method can be used for determining the predictability of some alarms. This can be a useful tool to identify possible alarms before the WT can be damaged.

5 Conclusions

In this work, a new methodology is presented for extracting information from SCADA dataset. The methodology is based on the generation of neural network from the quantitative (value of parameters) and qualitative dataset (alarms) of the SCADA system. The methodology has two different purposes. In first place, the adaptation of the SCADA control rules to the variable condition of the wind turbines. A neural network has been generated to determine which data should be added to the database. The creation of a "healthy" database allows for adapting the SCADA control rules to the real condition of the wind turbine over the time. Secondly, an additional neural network has been created for making predictions of the activation of alarms. This can be used to identify an abnormal state of the wind turbine earlier than the SCADA.

Acknowledgement. The work reported herewith has been financially by the Spanish Ministerio de Economía y Competitividad, under Research Grant Ref.: DPI2015-67264-P.

References

1. Arbib MA (2003) The handbook of brain theory and neural networks. MIT Press
2. Camacho E, Requena V et al (2014) Demonstration of methods and tools for the optimisation of operational reliability of large-scale industrial wind turbines. In: International conference on renewable energies offshore renew
3. Council GWE (2015) Global wind report annual market update 2014. Technical report, http://www.gwec.net/wp-content/uploads/2015/03/GWEC_Global_Wind_2014_Report_LR.pdf
4. Gómez Muñoz CQ, Márquez FPG (2016) A new fault location approach for acoustic emission techniques in wind turbines. Energies 9(1):40
5. González-Carrato RRDLH, Dimlaye V, Ruiz-Hernández D (2014) Pattern recognition by wavelet transforms using macro fibre composites transducers. Mech Syst Signal Process 48(1):339–350
6. González-Carrato RRDLH, Márquez FPG, Dimlaye V (2015) Maintenance management of wind turbines structures via mfcs and wavelet transforms. Renew Sustain Energy Rev 48:472–482
7. Kim K (2011) Use of scada data for failure detection in wind turbines. In: ASME 2011 5th International conference on energy sustainability, pp 2071–2079
8. Kusiak A, Li W (2011) The prediction and diagnosis of wind turbine faults. Renew Energy 36(1):16–23
9. Márquez FPG, García-Pardo IP (2010) Principal component analysis applied to filtered signals for maintenance management. Qual Reliab Eng Int 26(6):523–527
10. Márquez FPG, Muñoz JMC (2012) A pattern recognition and data analysis method for maintenance management. Int J Syst Sci 43(6):1014–1028
11. Márquez FPG, Pérez JMP, et al. (2016) Identification of critical components of wind turbines using fta over the time. Renew Energy 87
12. Marugán AP, Márquez FPG, Pérez JMP (2016) Optimal maintenance management of offshore wind farms. Energies 9(1):46
13. Masters T (1993) Practical neural network recipes in C++. Academic Press
14. Muñoz CQG, Márquez FPG, Tomás JMS (2016) Ice detection using thermal infrared radiometry on wind turbine blades. Measurement 93:157–163
15. Munoz CQG, Arenas JRT, Marquez FPG (2014) A novel approach to fault detection and diagnosis on wind turbines. Global Int J 16(6):1029–1037
16. Mylaraswamy D, Olson L, Nwadiogbu E (2007) Engine performance trending. In: AIAC12-HUMS conference
17. Song Z, Jiang Y, Zhang Z (2014) Short-term wind speed forecasting with markov-switching model. Appl Energy 130(5):103–112
18. Sun P, Li J et al (2016) A generalized model for wind turbine anomaly identification based on scada data. Appl Energy 168:550–567

19. Wymore ML, Dam JEV et al (2015) A survey of health monitoring systems for wind turbines. Renew Sustain Energy Rev 52:976–990
20. Yang W, Tavner PJ et al (2010) Cost-effective condition monitoring for wind turbines. IEEE Trans Industr Electron 57(1):263–271
21. Yang W, Court R, Jiang J (2013) Wind turbine condition monitoring by the approach of scada data analysis. Renew Energy 53(9):365–376
22. Zang X, Sun H, Trivedi KS (1998) A bdd-based algorithm for reliability graph analysis. Department of Electrical Engineering

Advances in Engineering Management of the Eleventh ICMSEM

Advances in Green Supply Chain, Resource Optimization Management, Risk Control and Integrated Project Management Based on the Eleventh ICMSEM Proceedings

Jiuping Xu(✉)

Uncertainty Decision-Making Laboratory, Sichuan University,
Chengdu 610065, People's Republic of China
xujiuping@scu.edu.cn

Abstract. Management Science and Engineering Management (MSEM) have significantly contributed to developments in society and economy, especially in management and control processes. In this paper, we first describe the basic concepts covered in the eleventh ICMSEM proceedings Volume II. We then conduct a review of engineering management (EM) research to identify the key areas, from which green supply chain, resource optimization management, risk control and integrated project management were found to be the most widely discussed research areas. After a summary of the key research achievements in the four areas, the high frequency EM keywords in the proceedings volume II are identified using NodeXL. The research trends both from MSEM journals and the ICMSEM are then summarized using the CiteSpace tool. As always, we are committed to providing an international forum for academic exchange and communication through the ICMSEM and plan to continue our innovative MSEM progress in the future.

Keywords: Green supply chain · Resource optimization management · Risk control · Integrated project management

1 Introduction

Management Science and Engineering Management, which emphasizes the theories, methods and engineering practice of complex management decision making, has been playing a vital role in both scientific endeavor and societal development. The increase in MSEM research in the past few decades has brought new vitality to management and engineering practice. At the same time, through the continuing and expanding development of innovative managerial tools, MSEM has significantly contributed to advancements in domestic and international economic development and increased general scientific management consciousness. MSEM is a complex synthesis with a broad research focus that combines complex management theories and practical engineering solutions to successfully solve management practice problems. This kind of cross-functional,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_76

multidisciplinary research supports real world management execution, improves management efficiency, and contributes to energy conservation.

MSEM research is widely applied to management problems that involve extensive engineering concepts. In proceedings volume I, the focus is on management science (MS) and its future development trends. Engineering management (EM) is therefore the main focus of proceedings volume II. Kocaolgu's defined EM as a field that examines engineering management using comprehensive scientific research to develop innovative engineering solutions that can improve technical organization, technical resources, and technical systems [5], in areas such as manufacturing, construction, design engineering, and industrial engineering. The ICMSEM proceedings Volume II focuses on four key EM areas; the green supply chain, resource optimization management, risk control, and integrated project management.

2 Literature Review

To better analyze the pertinent research fields and possible research directions, we reviewed the most popular research areas in the most recent EM research. What emerged was that the green supply chain, resource optimization management, risk control, and integrated project management have been the most widely studied in recent years. In this section, we review the related literature to analyze the developmental tracks in these four areas.

2.1 Green Supply Chain

The supply chain management concept emerged in the early 1980s, with the green supply chain (GSC) being first proposed by Michigan State University Manufacturing Research Association in 1996. Lambert et al. defined supply chain management as the integration of key business processes from the end user to the original suppliers to provide products, services, and information to add value for customers and other stakeholders [8]. Brazil Sarkis examined the social and cultural dimensions in green supply chain management and found that cultural boundaries govern the type of management skills and control processes used [7]. To empirically investigate the construction and scale of green supply chain management (GSCM) implementation by manufacturers, Zhu et al. tested two GSCM implementation measurement models and compared them using confirmatory factor analysis [19]. As a supply chain grows in scale and operation, its structure becomes more complicated. Some scholars have taken a fresh look at integrated green supply chain management [4, 13, 14] to ensure continuing sustainability.

2.2 Resource Optimization Management

Resource optimization management (ROM) is the efficient and effective deployment and allocation of an organization's resources when and where they are needed. Such resources may include financial resources, inventory, human skills, production resources, or information technology. Resources and energy shortages have become the main factors restricting the sustainable development of national economies, as low resource efficiency and high energy consumption can result in a serious waste of resources putting significant pressure on resources and environmental governance. Vadenbo presented a general multi-objective mixed-integer linear programming (MILP) optimization model aimed at providing decision support for waste and resource management in industrial networks [15]. Further, energy resource shortage problems associated with rapid social and economic development have been of critical concern to both national and local governments worldwide for many decades [9]. Water, as one of the most important resources on earth has therefore received a great deal of research attention, with water optimization management being a major focus in the past few years [1, 12, 18]. ROM involves all aspects of social life, so research strives to further develop this area using modern computer technology.

2.3 Risk Control

Risk control (RC) is the coordinated and economic application of resources to minimize, monitor and control the impact probability of unfortunate events, or to maximize the realization of opportunities after the identification, assessment and prioritization of risks. Muriana and Vizzinib presented a deterministic technique for assessing and preventing project risks by determining the risk of the Work Progress Status [10]. Valtonen et al. studied public risk management related to the use of public land development by analyzing case studies in Finland and the Netherlands, both of which have strong public land development traditions [16]. By identifying risks, specific state support and special project management measures have been developed to limit the negative influence of the possible project risks [17]. Therefore, developing a general framework to analyze corporate risk management policies and ensure risk control is vital [6]. In brief, RC includes transferring risk to another party by avoiding risk, reducing the negative effects of risk, and accepting some or all of the consequences of a particular risk.

2.4 Integrated Project Management

Integrated project management (IPM) is a philosophy that recognizes the different elements involved in projects to apply strong team leadership and encourage a collaborative ethos with clear purposes and strategies to ensure success. Planning, organizing, securing, and managing resources to successfully complete specific project goals and objectives are the main IPM research areas, the applications for which have been used in manufacturing, construction, design engineering, industrial engineering, technology, production and many other areas. Atkinson provided some thoughts about the success criteria for project management in which cost, time, and quality have become inextricably linked to project management success over the last 60 years [2]. An integrated methodology was developed for planning construction projects under uncertainty that relied on a computer supported risk management system to identify the risk factors in the integrated project [11]. IPM has also been applied to integrated waste management systems to identify the optimal breakdown between materials and energy recovery from municipal solid waste [3]. In all, integrated project management is a complex subject and needs to be examined from several perspectives.

3 Central Issues in the Proceedings Volume II

Keywords are an important index for the identification of essential content in research articles. The keywords in the 75 articles in Volume II were extracted to a database to allow for a tendency analysis. The keywords were sorted using the Keywords Match software that matched the keywords that were uniform or similar. Finally, the matched keywords were imported into NodeXL software to obtain the initial visual image shown in Fig. 1.

Fig. 1. DAS literature mining

With EM as the key component, the open source software tool NodeXL was used to facilitate the learning of the concepts and methods. To begin with, some key words were processed to unify expressions and keywords with the same meaning. This preliminary process reduced the number of keywords, making it possible to develop an efficient network. At the same time, the vertices' shapes were set to determine the betweenness and closeness centrality. When the degree of a vertex's betweenness and closeness centrality was more than five, the shape of this vertex was a red diamond in the resulting figure. The aim of this analysis was to determine the key EM concepts that connected with other research topics through the primary nodes. After calculating, clustering and filtrating, the results in Fig. 2 illustrate the important research relationships.

The above analysis highlighted the key areas focused on in the proceedings volume II; green supply chain management, resource optimization management, risk control and integrated project management problems related to green and pro-environment concepts. Tao focuses on a simplified manufacturing and distribution supply chain planning network focused on carbon emission constraints, for which a bilevel programming model was mathematically formulated. Through the study of related emissions policies, Lu finds that under different carbon policies, revenue sharing contracts are able to coordinate the supply chain with the manufacturer's revenue sharing proportion under

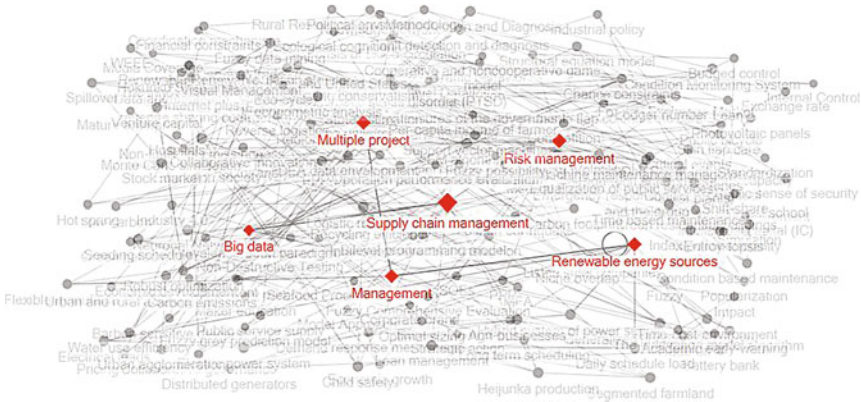


Fig. 2. Research topics in EM for the Eleventh ICMSEM

a cap-and-trade being always less than a situation in which there is no carbon emissions constraint. Machado and Duarte present an overview of Industry 4.0, a new developing concept in the field of industrial engineering, and lean and green supply chain management, for which a conceptual model is developed. Wang et al. concludes that when big data suppliers are part of the supply chain competition and one party gains a dominant supply chain position, collaborative decision-making is the key to enhancing overall supply chain profitability. These studies highlight the many new characteristics involved in green supply chains, such as big data and carbon emissions reductions.

Resource optimization management primarily covers resource processing, exploitation, production, and consumption. Arifjanov and Zakhidov consider an issue for the development of methods for the estimation of electrical loads for various electricity customers in residential and public buildings, and address issues related to the integration of distributed generators to the UES to ensure the optimal management of technological modes. Gómez Munoz et al. develops an optimal maintenance strategy that includes the NDT system to reduce costs and increase the competitiveness of renewable energy sources. For water resource optimization, Tu et al. proposes an equity and sustainability based model that considers equity, economic benefit, and ecological environment and Liu et al. uses a DEA method to establish an evaluation model to empirically study the urban agglomeration of Chengdu city from 2008–2014. Because of the wide use and serious shortage of some resources around the world, scientific research into resource optimization management has increased to assist governments and enterprises around the world.

Risk control is important in all organizations and enterprises. In this section, Zhou et al. presents a model to mitigate the risk of budget overruns during raw materials procurement, and Chen et al. provides a risk allocation model using a fuzzy comprehensive evaluation method aimed at optimizing risk allocations, reducing transaction costs, and enhancing comprehensive benefits from the perspective of state-owned enterprises. Beside these studies, to better control risks occurring in financial markets, Shibli et al. investigates the link between the foreign exchange markets and the stock market in

Bangladesh by examining the volatility spillover between the markets, volatility persistence, and the asymmetric effect of information on the volatility of these two financial markets. In addition, Zhang et al. studies a projection pursuit risk assessment model using a combined method to model PPP risk under the background of Big Data. Effective risk control can reduce potential risk factors and assist managers gain increased benefits.

The last section in Volume II focuses on the developments in integrated project management. Elchan defines the scientific problems associated with management at the beginning stages of an innovation project in the departments of a technology park at a higher education school in Azerbaijan. Based on principal-agent theory and game theory, a theoretical framework for a knowledge network conflict coordination mechanism is constructed by Wei, which divides the conflict coordination mechanism into three levels; contract mechanism, self-implementation mechanism, and a third-party conflict coordination mechanism. Sheng et al. proposes suggestions and countermeasures for the equalization of public services in urban and rural areas from three integrated project aspects and Tu et al. develops two staged fuzzy DEA models with undesirable outputs to evaluate the banking system. The overall process of leading, organizing, staffing, planning, and controlling activities requires managers to use systems viewpoints, methods, and theories to optimize the work involved when seeking effective integrated project management under limited resource constraints.

4 Evaluation of EM and ICMSEM Development Trends

In this section, we evaluate the ICMSEM to validate that the research trends are in line with the trends emerging in current MSEM journals. At first, CiteSpace, an information visualization technology developed by Chen, was utilized to draw the a scientific knowledge map that display the development trends in a discipline or knowledge domain over a certain period, and identifies the possible evolution through an analysis of the research frontiers.

"Advanced search" with "engineering management" was used as the identifier in the "web of science" database with a timespan from 2000 to 2017. Initially, 42499 published articles were identified and after the final screening using TI = (engineering management) 2963 articles were reviewed.

4.1 The Development Trends of EM

The statistics from the 2963 articles recorded output were saved and converted into CiteSpace which transformed the data into a format that could be identified by the software to allow for parameter selection. In this operation, the time span was from 2000 to 2017 with the time slice set at one year and the theme selection based on the titles, abstract subject words, identifiers, and keywords to allow for node selection. Then, each zone with the 30 highest keyword records were clustered and analyzed, from which a map was drawn for the minimum spanning tree. As shown in Fig. 3, by setting the

"Threshold = 30", a total of 348 nodes were obtained, with the overall network density being 0.0098. Therefore, the system frequency identified green supply chain management, environmental management, project management, risk management, and models and systems as the highest ranked areas. This not only displayed the most popular research fields in engineering management but also implied the future EM development trends.

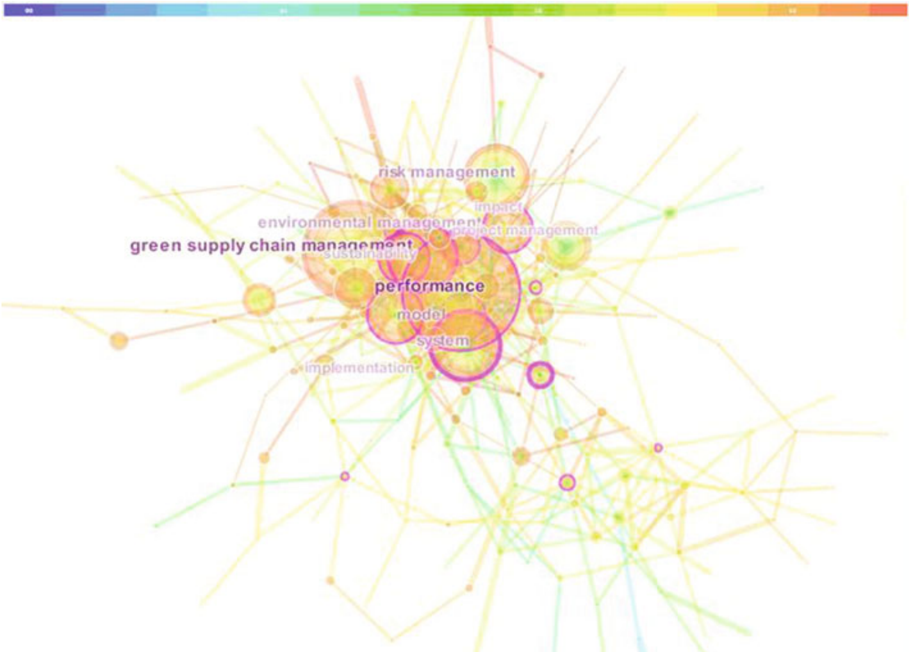


Fig. 3. The results of co-occurring keywords in EM

With a reference frequency running from high to low, the top thirty were analyzed from the 348 keywords. As shown in Table 1, the keywords such as green supply chain management, risk management, environmental performance and sustainability had a relatively high centrality.

Using the keyword with the label title clustering, 40 categories were identified; however, only the top 10 categories are shown in Fig. 4. The other topics were relatively dispersed so are not displayed, which indicated that scientific engineering management research was relatively loose. In particular, the new keywords for 2015 and 2016 have not resulted in significant research attention.

Table 1. The top thirty hot keywords of EM

Frequency	Centrality	Year	Cited references
84	0.11	2010	Performance
77	0.03	2011	Green supply chain management
54	0.07	2002	Model
54	0.08	2002	Risk management
52	0.27	2007	System
48	0.21	2011	Environmental management
41	0.15	2007	Implementation
40	0.1	2012	Sustainability
39	0.05	2007	Project management
38	0.12	2011	Impact
36	0.11	2011	Industry
32	0.05	2011	Strategy
29	0.08	2010	Perspective
29	0.12	2010	China
25	0.04	2011	Green supply chain
24	0.05	2010	Framework
21	0.06	2011	Innovation
21	0.04	2009	Integration
20	0.05	2008	Firm
20	0.05	2011	Reverse logistics
20	0.05	2011	Design
17	0.01	2012	Capability
17	0.02	2013	Initiative
16	0.21	2004	Risk
16	0.05	2011	Pressure
16	0.01	2010	Environmental performance
15	0.1	2008	Quality
15	0.01	2013	Selection
15	0.06	2007	Uncertainty
15	0.05	2008	Optimization

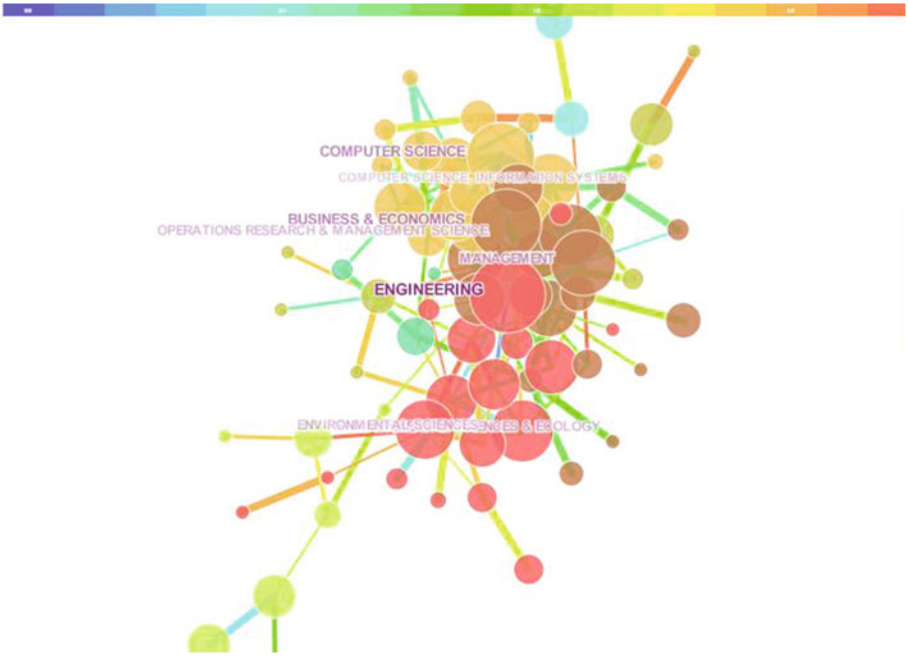


Fig. 4. The results of research fields clustering of EM

4.2 Future Development Predictions

To further understand the focused engineering management research areas, we examined each category to determine the common themes. At "Threshold = 30", the high fre-

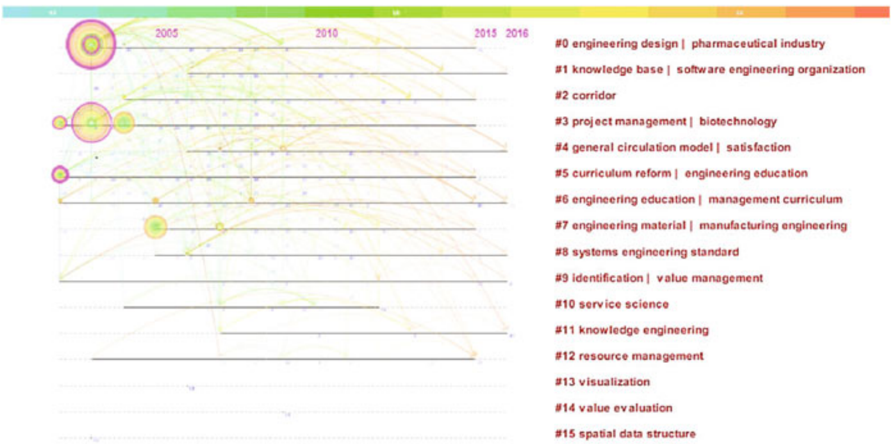


Fig. 5. The category clustering of EM timeline view

Table 2. The results of top thirty categories

Frequency	Centrality	Year	Cited references
307	0.21	2002	Engineering
185	0.01	2004	Business & Economics
176	0.01	2002	Computer Science
136	0.13	2005	Management
104	0.05	2002	Operations Research & Management Science
102	0.02	2002	Environmental Sciences & Ecology
95	0.22	2002	Environmental Sciences
73	0.12	2002	Computer Science, Information Systems
71	0.07	2002	Engineering, Electrical & Electronic
70	0.02	2005	Engineering, Industrial
68	0.13	2004	Business
60	0.02	2007	Engineering, Civil
51	0.17	2006	Economics
50	0.02	2007	Computer Science, Theory & Methods
50	0.1	2005	Computer Science, Interdisciplinary Applications
50	0.09	2002	Engineering, Environmental
47	0.23	2004	Water Resources
42	0.05	2006	Business, Finance
41	0.12	2010	Science & Technology - Other Topics
39	0	2002	Cardiovascular System & Cardiology
39	0.04	2007	Engineering, Manufacturing
38	0.06	2006	Computer Science, Artificial Intelligence
32	0	2013	Green & Sustainable Science & Technology
28	0.11	2005	Construction & Building Technology
28	0.04	2002	Telecommunications
27	0	2004	General & Internal Medicine
26	0.08	2004	Medicine, General & Internal
25	0	2002	Cardiac & Cardiovascular Systems
24	0.04	2010	Materials Science
23	0	2003	Peripheral Vascular Disease

quency research areas and timezone view diagram were identified, as shown in Fig. 5. Most of these high frequency words appeared in the early days, indicating that engineering management research has been focused around these topics for quite a long time. The analysis of the top 30 most frequent words (Table 2), found that engineering, business and economics, computer science, environmental science and operations research are currently the most popular research areas.

From Tables 1 and 2, the papers presented in this year's ICMSEM proceedings volume II closely reflect the most pertinent engineering management research areas; green supply chain (systems management, industry, green and environmental protection), resource optimization management (environmental sciences, water resources, engineering education), integrated project management (project management, industry, integrated project) and risk control (risk management, risk, design, information). In addition, the computer science and environmental science research highlights how high-tech can spur social progress and environmental protection.

We believe that EM should focus on the study of specific EM problems as well as popularizing MS knowledge. Excellent academic research can effect developments across the world, but a further focus on regional areas is also needed. To ensure a bright EM future, there is a need for inspiration, practical theories, effective methods, and extensive applications in future developments. To achieve this, EM knowledge needs to be popularized, which is the duty of all MSEM academics. In the future, more focus on low carbon emissions, environmental protection, big data, energy utilization, and other popular EM issues are needed.

5 Conclusion

Engineering management is a complex area that involves all engineering aspects. The open source software tool NodeXL identified the four areas covered in the eleventh ICMSEM proceedings Volume II, from which we summarized the key research in this year. To analyze the EM and ICMSEM development trends, we identified the main search terms and keywords using CiteSpace. Our key objective is to continue to improve the quality of papers in the proceedings and to ensure the ICMSEM organization is dynamic and appealing to EM researchers worldwide. EM research is continuously developing and new trends are appearing every year; however, more research is necessary so as to popularize EM developments and provide a more active research forum.

Acknowledgements. The author gratefully acknowledges Jingqi Dai and Lin Zhong's efforts on the paper collection and classification, Zongmin Li and Lurong Fan's efforts on data collation and analysis, and Ning Ma and Yan Wang's efforts on the chart drawing.

References

1. Afshar A, Massoumi F et al (2015) State of the art review of ant colony optimization applications in water resource management. Water Resour Manage 29(11):3891–3904
2. Atkinson R (1999) Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. Int J Project Manage 17(6):337–342
3. Consonni S, Giugliano M et al (2011) Material and energy recovery in integrated waste management systems: project overview and main results. Waste Manage 31(9–10):2057–2065
4. Fahimnia B, Sarkis J, Davarzani H (2015) Green supply chain management: a review and bibliometric analysis. Int J Prod Econ 162:101–114
5. Farr JV, Buede DM (2003) Systems engineering and engineering management: keys to the efficient development of products and services. Eng Manage J 15(3):3–9

6. Froot KA, Scharfstein DS, Stein JC (1993) Risk management: coordinating corporate investment and financing policies. J Financ 48(5):1629–1658
7. Govindan K, Sarkis J et al (2014) Eco-efficiency based green supply chain management: current status and opportunities. Eur J Oper Res 233(2):293–298
8. Lambert DM, Cooper MC (2000) Issues in supply chain management. Ind Mark Manage 29(1):65–83
9. Ming Z, Song X et al (2013) New energy bases and sustainable development in China: a review. Renew Sustain Energ Rev 20(4):169–185
10. Muriana C, Vizzini G (2017) Project risk management: a deterministic quantitative technique for assessment and mitigation. Int J Project Manage 35:320–340
11. Schatteman D, Herroelen W et al (2008) Methodology for integrated risk management and proactive scheduling of construction projects. J Constr Eng Manage 134(11):885–893
12. Singh A (2014) Simulation and optimization modeling for the management of groundwater resources. I: distinct applications. J Irrig Drainage Eng 140(4):04013,021
13. Singh A, Trivedi A (2016) Sustainable green supply chain management: trends and current practices. Competitiveness Rev 26(3):265–288
14. Srivastava SK (2007) Green supply-chain management: a state-of-the-art literature review. Int J Manage Rev 9(1):53–80
15. Vadenbo C, Guilléngosálbez G et al (2014) Multi-objective optimization of waste and resource management in industrial networks - part ii: model application to the treatment of sewage sludge. Resour Conserv Recycl 89(4):52–63
16. Valtonen E, Falkenbach H, Krabben EVD (2017) Risk management in public land development projects: comparative case study in Finland, and The Netherlands. Int J Covering Aspects Land Use, Land Use Policy 62:246–257
17. Wang HB (2007) Application of risk management to wind power project. Sci Technol Manage 43:48–51
18. Wang R, Li Y, Tan Q (2015) A review of inexact optimization modeling and its application to integrated water resources management. Front Earth Sci 9(1):51–64
19. Zhu Q, Sarkis J, Lai KH (2008) Confirmation of a measurement model for green supply chain management practices implementation. Int J Prod Econ 111(2):261–273

Green Supply Chain

Modelling a Supply Chain Network of Processed Seafood to Meet Diverse Demands by Multi-branch Production System

Koichi Murata[1(✉)], Seiichiro Isobe[1], and Hiroshi Katayama[2]

[1] College of Industrial Technology, Nihon University, Tokyo, Japan
murata.kouichi30@nihon-u.ac.jp
[2] Faculty of Science and Engineering, Waseda University, Tokyo, Japan

Abstract. The demand for processed seafood has recently increased owing to diversity of taste and the crises of food supplies. Processed seafood companies have the potential to manufacture various products by combining fish varieties and processing methods. They also associate with many customers such as other food processing companies, food service industries, FMCG companies, and end users. This study aims to reveal the supply chain model of processed seafood to multi-customers by the multi-branch production system. A model is constructed to understand the practical system from a theoretical perspective. The main finding of the study is that four paths of the proposed network model between a seafood processing company and customers exist: (1) defrosting to processing, (2) cleaning to service, (3) processing to service, and (4) seasoning to retail. These findings clarify the supply chain network in the industry and help consider business opportunities.

Keywords: Food supply chain management · Seafood processing industry · Model approach

1 Introduction

Fish and seafood have recently become important materials to meet diversified customer needs resulting from the food crisis caused by population explosion and the requirement of seafood which is essentially healthy (i.e. low calorie count and nutrient intakes) and which provides convenient meals for a two-income family, a child-rearing family, and aging society in advanced counties. To match the resultant explosive demand, the globalization of the seafood supply chain has increasingly progressed with technological innovation. For example, the cold chain has expanded to maintain the freshness of materials through improvement of freezing technologies for transport worldwide [11]. A traceability system has been developed to obtain the trust of customers by informing them of the safety of materials by a physical sensor system and information technologies [6].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_77

In the Japanese context, to aim for the creation of a tourism nation and for the next summer Olympic Games in 2020, an enrichment of the food supply chain is required. The Japanese dietary culture called 'Washoku', registered as a 2013 Intangible Cultural Heritage by UNESCO, must be a powerful weapon for executing the strategy. Processed seafood, one core ingredient of 'Washoku', is manufactured in small and medium enterprises (SMEs) mostly located in the area that experienced the Great East Japan earthquake (Higashi Nihon Daishinsai) on 11 March 2015. Furthermore, the occurrence of climate change including global warming and abnormal weather and over-fishing by neighbouring countries causes a decrease of marine resources in the Pacific Ocean. These SMEs operate on a small scale and sell processed seafood to identified customers according to their requirements. For the future, they must overcome difficulties and reinvent themselves to obtain new markets and customers worldwide through business innovation.

Many studies related to the seafood supply chain have presented new perspectives, including open innovation of seafood value chain [12], an international distribution system [1,2], firm structure [4], quality assurance with labelled seafood products [7], a sustainable system [3], marketing and economic innovation [5,9,14], seafood supply chain management [10], and an inventory system [8]. However, there are not enough studies related to innovation of the business model which give an overall perspective to change the present business structure.

Based on the recognition of current practical and academic conditions, the present study aims to explore the business model of the processed seafood industry with the concept level as the starting point of the study. Specifically, the network between a food processing company and its customers is focused on understanding the exciting business style of the food processing company and grasping the possibility of cultivating routes to new customers.

2 Methodology

This study tries to draw the network structure between one food processing company and its candidate customers because a clarification of the relation among participants firstly needs to retrieve investment points for business expansion. The description is performed based on an observation and an interview of several seafood processing companies in the rebuilding project after the Great East Japan earthquake for three years. The authors of the present study have gradually understood these companies business structure through investigation and a trial-and-error method. This study presents the results. The method to build the described model is business modelling which reveals the following things pursuant to the aim of the study [15]:

The business model is emerging as a new unit of analysis.

- Business models emphasize a system-level, holistic approach to explain how firms 'do business'.
- Firm activities play an important role in the various conceptualizations.

- Business models seek to explain how value is created, not just how it is captured.

This research also reference typical models related to a production system, such as make-to-order (MTO) and make-to-stock (MTS) models [13].

Four steps are adopted in this study. The first is based on the seafood cooking process. Seafood processing companies cook materials at the location of mass customers. The step will be useful to identify fundamental processes in a production system for seafood processing. The second is the deployment of candidates of a customer. They are other food processing companies, food service companies, fast-moving consumer goods (FMCG) companies, and consumers. This step describes the details of the four customers because the purpose of the step is to become aware of new customers that differ from existing customers. In the third step, a network model is constructed to visualize a possibility of new business through a combination of outputs in the first and second steps. The final step is a confirmation of the capability of the constructed network through an initial discussion of business expansion in the future. The flowchart of all steps is provided in Fig. 1.

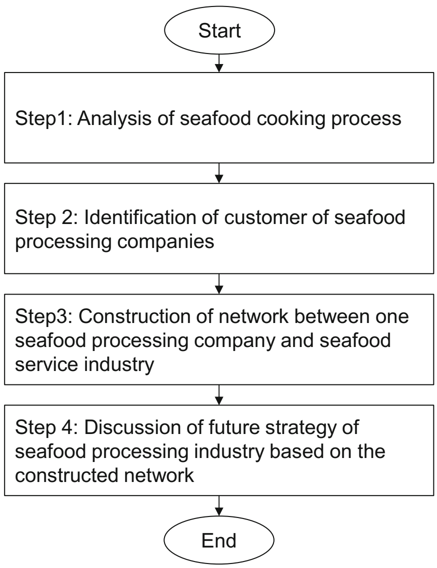


Fig. 1. Research procedure.

3 Construction of Seafood Supply Chain Model

This section describes the result of each step illustrated in Fig. 1.

3.1 Seafood Cooking Process (Step 1)

The journey to develop a processed seafood business model in this study starts with the understanding of daily eating habits as follows. The main processes to eat seafood have not changed from ages past, as shown in Fig. 2: fishing, cleaning, processing, seasoning, and eating. The middle three processes exist mainly to add value to fresh fish. These can be omitted and exchanged in the cooking process. Many techniques and skills are included to make the materials taste even better. The detailed explanation of each type is as follows:

(1) Fishing process

This process is considered a preparation part for processed seafood production involving obtaining the materials. The quality level of the materials depends on size, weight, appearance, freshness, variety, and so forth. Over-fishing and climate change seriously affect a haul of fish. A fishing quota by multiple nations and enclosed aquaculture with the newest biotechnologies are countermeasures for sustainable marine resources.

(2) Cleaning process

The complicated body of fish and seafood causes a decline in the productivity of seafood processing. It is basic and important to protect customers' safety from dangerous parts which they cannot eat (i.e. hard fish, thick bone, and the internal organs of a fish including poisonous substances). The variety of processes includes carving, boning, scaling, and filleting. Accurate fileting of large fish like salmon requires immense skill with yield rate and waste disposal.

(3) Processing process

This process is the main portion of seafood processing. Representative methods of the process are grilling, reducing, frying, steaming, and drying. The selection of each method relates to a core business of a seafood processing company with a large investment in production facilities. The design of the process profile of production by the selected method is a company secret relating to the organoleptic feel of processed seafood because many companies adopt the common categories of methods noted above.

(4) Seasoning process

This process is the portion that adds tastiness and flavour-value to a processed item. In Japan, there are five traditional popular seasonings, such as sugar (Sato), salt (Shio), vinegar (Su), soy sauce (Shoyu), and soybean paste (Miso). They characterize the taste of Japanese seafood called 'Washoku', which was registered as a 2013 Intangible Cultural Heritage by UNESCO. Chemical seasoning widens the possibility of new ways to enjoy processed seafood.

(5) Eating process

This is the final process to synthetically evaluate the quality of the four processes by consumers. In this stage, not only the taste of processed seafood, but also a dishing, including a selected plate and a side dish, influences the total impression. The family styles of advanced countries, such as a two-income family, a child-rearing family, and the aging society, also require the convenience of a microwave oven and reheat pouches.

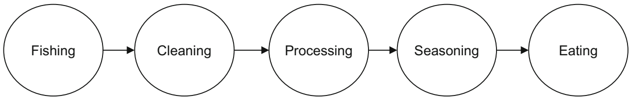


Fig. 2. Basic process to eat seafood

In the present age, the division of the five processes between fisheries and consumers depends on consumers' requirements. It indicates the following five models proposed in Fig. 3.

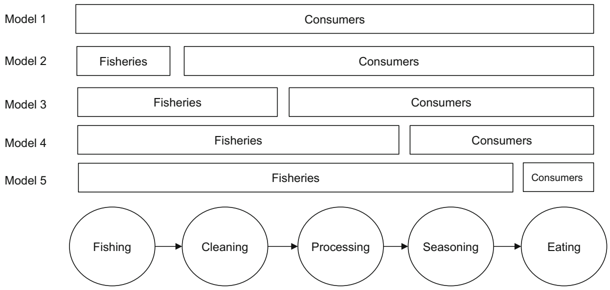


Fig. 3. Division of roles between fisheries and consumers

Consumers do not need support from fisheries in Model 1, which constitutes a hobby. In the case of Model 2, the fishing industry will exist in a market near a fishing port with consumers. In the three models consisting of two processes and more by fisheries (i.e. Models 3 to 5), the division of roles occurs among three industries: (I) fishing industry, (II) seafood processing industry, and (III) seafood service industry including restaurant and retail businesses. The division is considered based on the two following assumptions:

- The fishing industry is necessarily in charge of the fishing process.
- The orders of the three industries do not become reversed.

Nineteen sub-models of the three models are summarized in Table 1. Models 3, 4, and 5 contain 3, 6, and 10 sub-models. Processed seafood is sold directly to consumers or through a mail-order system and internet service when the sub-model consists of (I) the fishing industry and (III) the seafood service industry. (II) The seafood processing industry asks (III) the seafood service industry to sell their products in addition to the above-mentioned selling methods when these two entities exist in the sub-model. This table shows that (II) the seafood processing industry can choose, in theory, appropriate sub-models from the alternatives. The flexibility seems to suggest the possibility of expanding the business of each seafood processing company.

Table 1. Sub-models of Models 3 to 5

Model	Fishing	Cleaning	Processing	Seasoning	Note
3	I	I	-	-	Selling directly to customers
	I	II	-	-	Selling by a service industry or directly to customers
4	I	I	I	-	Selling directly to customers
	I	I	II	-	Selling by a service industry or directly to customers
	I	I	III	-	Selling directly to customers
	I	II	II	-	Selling by a service industry or directly to customers
	I	II	III	-	Selling directly to customers
	I	III	III	-	Selling directly to customers
5	I	I	I	I	Selling directly to customers
	I	I	I	II	Selling by a service industry or directly to customers
	I	I	I	III	Selling directly to customers
	I	I	II	II	Selling by a service industry or directly to customers
	I	I	II	III	Selling directly to customers
	I	I	III	III	Selling directly to customers
	I	II	II	II	Selling by a service industry or directly to customers
	I	II	II	III	Selling directly to customers
	I	II	III	III	Selling directly to customers
	I	III	III	III	Selling directly to customers

(I) fishing industry, (II) seafood processing industry, and (III) seafood service industry.

3.2 Customers of Processed Seafood Industry (Step 2)

Customer candidates for (II) the processed seafood industry are rich in variety. The main candidates are four types of seafood service companies: other food processing companies, food service companies, fast-moving consumer goods (FMCG) companies, and consumers. The explanations of each type are as follows:

(1) Other food processing companies

These customers use the product supplied from seafood processing companies as one material of their final products. For example, raw materials are supplied from a company which manages a large-capacity frozen warehouse to store them cheaply and in huge quantities. Cleaned materials (i.e. seafood paste and fish cut into small cubes) also have high value for customers who do not have technical know-how.

(2) Food service companies

These customers are food service companies including restaurant businesses; home-delivery services; services for the provision of meals by a school, hospital, or welfare facilities. They basically have a kitchen which has the capacity for making and serving meals with materials supplied from seafood processing companies to their customers. Moreover, if their business model is a franchise chain, they construct a supply network with a central factory which executes the middle processing of materials for speedy and effective servicing of meals at all their shops.

(3) FMCG companies

These customers are types of FMCG companies (e.g. a convenience store, supermarket, shopping mall, or department store). They mainly offer two services. One is the direct sale of goods supplied from seafood processing companies. The other is the sale of a daily dish including simple processed/cooked materials. Their customers are end users who live near these small shops. FMCG companies obtain large sales accumulated across their many branches, but the number of end users is limited for one store.

(4) Consumers

Seafood processing companies usually do not directly access end users except in face-to-face sales like at a regional festival and special sales once a year. They contact end users to sell their products through a mail-order system with mass-communication and a system of online electronic commerce. This opportunity will increase in the era of the internet.

3.3 Network Between Seafood Processing Company and Seafood Service Industry (Step 3)

Figure 4 illustrates the relationship between one seafood processing company and seafood service companies. The left side of the figure is a typical production system in a seafood processing company. This system mainly consists of four processes: defrosting, cleaning, processing, and seasoning. The first one is a process to prepare frozen materials and deliver them to maintain their freshness while restraining freezer burn if possible, by a fishing company, and to maintain a sufficient inventory. The other three processes add appropriate values to the defrosted materials described in the previous section. The notable characteristic of the system is a multi-branch production system with the potential to supply semi-processed products as final products to customers.

Owing to the capability of the industry, candidate customers are rich in variety, as shown in the right side of the figure. Main candidates are the four types of seafood service companies mentioned above (i.e. other food processing companies, food service companies, FMCG companies, and consumers). While centring on seafood processing companies, they should consider all combinations between processes of seafood processing and four types of seafood service companies. In particular, the following four paths are useful for food service companies to realize their service while reducing initial workloads.

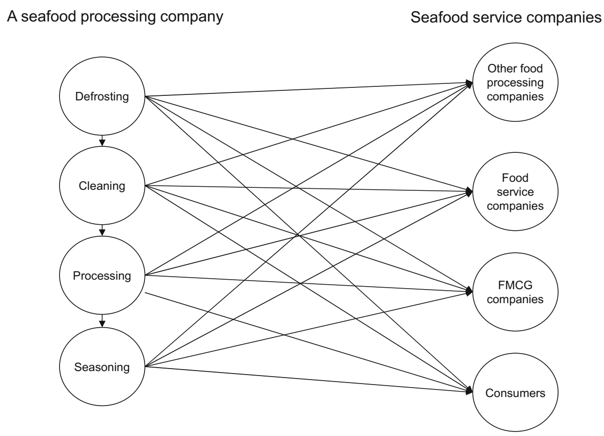


Fig. 4. Relationship between a seafood processing company and seafood service companies.

(1) A path from defrosting to processing (DTP)

In this path, seafood processing companies supply frozen/defrosted materials to other food processing companies. The added-value is a supply system of materials sufficient both in quality and quantity. Seafood processing companies must improve the refrigeration capacities of inventory and transportation. Owing to these functions, other food processing companies outsource procurement operations.

(2) A path from cleaning to service (CTS)

In this path, seafood processing companies supply frozen materials to seafood service companies. The added-value is a system to clean masses of fish and seafood. The seafood processing company must improve its capability through training of skills related to the operation and its automation. Owing to this function, food service companies will concentrate on creative cooking and close servicing, which are considered to be their core functions.

(3) A path of processing to service (PTS)

In this path, seafood processing companies supply processed materials to seafood service companies. The added-value is a mass production system of processed seafood. Seafood processing companies should improve the main process in their business. Owing to this function, food service companies will provide effective and uniform services to their customers with simple cooking.

(4) A path of seasoning to retail model (STR)

In this path, seafood processing companies supply seasoned materials to FMCG companies. The added-value is a mass supply system of processed seafood. Seafood processing companies should improve the main processes in their business. Owing to this function, retail companies can sell various moderately priced and tasty items to end users every day.

3.4 Future Strategy for Seafood Processing Company (Step 4)

To confirm the utility of the constructed network model in the previous sections, a future strategy of seafood processing companies will be considered in this section.

Most SMEs continue to do business along one path of the network model described in the proposed network in Fig. 4. They may consider the path as the only possibility for their business with limited management resources. However, they should understand two characteristics of their industry.

One is the flexibility of their production system. It is possible to supply products from any process in their production system. For example, if the current path of the network model is STR, semi-products of the path become fully final products for the other three paths (i.e. DTP, CTS, and PTS) without large facility investment. The other is that they have many candidate customers. If the current path of the network model is CTS, the same product could be supplied by PTS. The development of a new path is naturally difficult, even if the potential of the path is recognized. However, it will be a valuable trial to expand business in the future.

The utility of the proposed network is confirmed because it is useful to consider a subject to realize new business possibilities with the characteristic production system based on the above discussion.

4 Concluding Remarks

This study aims to reveal the supply chain model of processed seafood to multiple customers via a multi-branch production system. The findings consist of four paths: (1) DTP, (2) CTS, (3) PTS, and (4) STR. They clarify the supply chain network in the industry and consider the opportunity of the business. A future study could provide a mathematical formulation of the network model to quantitatively simulate and evaluate the impact of each path.

References

1. Abrahamsen MH, Håkansson H, Naudè P (2007) Perceptions on change in business networks: Norwegian salmon exporters and japanese importers. In: 23rd Annual IMP Conference, Manchester
2. Abrahamsen MH, Naudé P, Henneberg SC (2011) Network change as a battle of ideas? analysing the interplay between idea structures and activated structures. Imp Group 5(2):122–139
3. Alden R (2011) Building a sustainable seafood system for maine. Maine Policy Rev 20(1):87–95
4. Brydon K, Dana LP (2011) Globalisation and firm structure: comparing a family-business and a corporate block holder in the new zealand seafood industry. Int J Globalisation Small Bus 4(2):206–220
5. Chang KS, Waters EC (2006) The role of the alaska seafood industry: a social accounting matrix (SAM) model approach to economic base analysis. Ann Reg Sci 40(2):335–350

6. Farooq U, Tao W et al (2016) epedigree traceability system for the agricultural food supply chain to ensure consumer health. Sustainability 8(9):839
7. Jaffry S, Pickering H et al (2004) Consumer choices for quality and sustainability labelled seafood products in the UK. Food Policy 29(3):215–228
8. Murata K, Watanabe N et al (2015) Diagnosis of production system of marine frozen products by inventory management theory - a case of blue fins. Int J Intell Inf Syst 4(2–1):18–24
9. Quagrainie KK (2006) Analysis of US catfish fillet market share using a flexible logistic model. Marine Res Econ 21(1):33–45
10. Roheim CA (2008) Seafood supply chain management: Methods to prevent illegally-caught product entry into the marketplace. IUCN World Conservation Union
11. Salin V, Nayga RM (2003) A cold chain network for food exports to developing countries. Int J Phys Distrib Logistics Manag 33(10):918–933
12. Silva DD, Bjondal T (2013) An open innovation and its role in global fish and seafood value chains: Beyond the conventional wisdom
13. Vollmann B, Berry W, Claywhybark D (1997) Manufacturing Planning and Control Systems. Irwin, New York
14. Wessells CR, Anderson JL (1992) Innovations and progress in seafood demand and market analysis. Mar Resour Econ 07(4):209–228
15. Zott C, Amit R, Massa L (2011) The business model: recent developments and future research. J Manag 37(4):1019–1042 Official Journal of the Southern Management Association

Hybrid Global Supply Chain Design Considering Product Families and Layered Cellular Design

Jue Jiang and Gürsel A. Süer(✉)

Department of Industrial and Systems Engineering,
Ohio University, Athens, OH 45701, USA
suer@ohio.edu

Abstract. This paper proposes a new hybrid global supply chain strategy for a global blood sugar manufacturing company. In this approach, products are analyzed based on product families. Product families can share cells across the globe if needed. The results are compared with the host market strategy where three manufacturing plants located in three regions to meet world demand. Manufacturing plants are designed considering layered cellular design approach under stochastic demand. This approach allows three cell types; (1) dedicated cells where each dedicated cell can be used only by one family, (2) shared cells where a cell can be shared by two product families, and (3) remainder cells where a cell can be used by three or more families. The main focus of this paper is to compare both alternatives with respect to number of manufacturing cells and number of machines needed. In this case, number of machines remained the same for both approaches. However, the proposed new approach led to more dedicated cells.

Keywords: Global supply chain · Layered cellular manufacturing · Stochastic demand

1 Introduction

This research aims designing a manufacturing system for a global blood sugar strip manufacturer. Three manufacturing facilities located in different continents are assumed to meet the demand of these regions. By using layered cellular manufacturing concepts, number and type of manufacturing cells are determined for each manufacturing facility considering stochastic demand data and production rates. Later, this supply chain strategy is compared with the proposed one in this paper where some cells are shared globally. A probabilistic method is used to design the cells.

The type of a manufacturing system mostly depends on the layout of the manufacturing system. As mentioned in Süer, Huang and Maddisetty [19], manufacturing systems can be classified into four categories based on the layout: process layout, fixed layout, cellular layout and product layout. Fixed Layout

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_78

is used when products are heavy and large and therefore are mostly stationary. Therefore, products typically stay in the same position and workers, machines and equipment are brought to the products to perform the tasks [19]. Product Layout is appropriate when production volume is high and product variety is low. Product layout is usually very efficient since it can be configured to perfection for a single product but this makes it an inflexible system. Process Layout is the best system for low production volume systems with high product variety [19]. These systems are very flexible since they can handle many different products but unfortunately they are not very efficient since they requires extensive material handling and thus lead to long leadtimes. On the other hand, Cellular Layout is more flexible than Product Layout and more efficient than Process Layout. It suits well for high product variety with low to moderate demand [19]. Figure 1 shows the classification of four layout types. Obviously, a manufacturing system can indeed consist of any number of layout types as well such both Cellular Layout and Process Layout etc.

Cellular Manufacturing is based on the grouping of similar products with respect to similar processes into one or more similar cells as required. In the real manufacturing world, many uncertainties exist in the system such as demand uncertainty, supply uncertainty and processing uncertainty. These uncertainties have been discussed extensively in many research works. Süer, Huang and Maddisetty [19] proposed to consider the uncertainties of product demand and processing times. By probabilistic market demand calculation, the part-family assignment is achieved [19]. Then, low utilized cells are grouped to increase the utilization of the system i.e., to reduce the number of machines and thus investment, even though some of these cells may no longer be pure cells. These different cell types, namely, dedicated (DC), shared (SC) and remainder (RC) cells are also shown in Fig. 1.

Supply chain is the network connecting suppliers, manufacturers, distribution centers and customers [16]. Dicken [7] discussed many supply chain models.

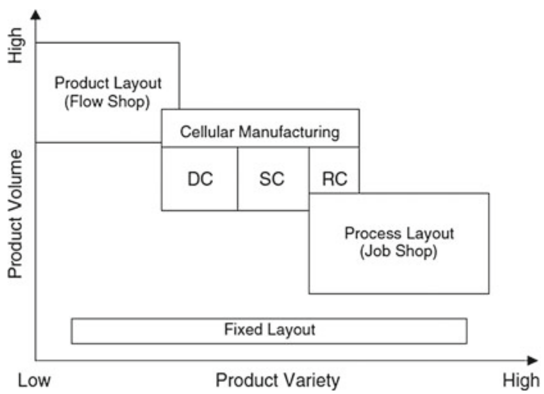


Fig. 1. Four types of manufacturing layout [19]

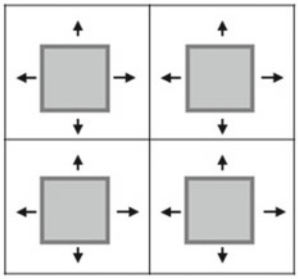


Fig. 2. Host market production model [7]

Among them, host market production model is illustrated in Fig. 2. In this strategy, each of the geographic regions covers its own demand. In this study, this strategy will be compared with the newly proposed hybrid global supply chain strategy that considers product family definitions and layered cellular approach.

2 Literature Review

Literature regarding global supply chain and cellular manufacturing systems is summarized in this section.

Chopra and Meindl [6] introduced that supply chain model includes three logistical drivers and three cross functional drivers. They defined logical drivers as facilities, inventory and transportation and three cross functional drivers as information, sourcing and pricing. Brush, Maritan and Karnani [2] mentioned that many factors affected a network of multinational manufacturing facilities. Besides considering facility selection, Haug [9] also considered production shifting. A facility location model was developed to study the location decision of high technology firms. The model identified the international manufacturing facility location based on domestic and potential international production markets. This allowed production to be transferred from domestic manufacturing facilities to foreign facilities. This allowed production to be transferred from domestic manufacturing facilities to foreign facilities. Lowe, Wendell and Hu [13] included production capacity to develop a two-phase multi-screening approach to incorporate uncertainty about exchange rates and exchange rate risk. Canel and Das [4] improved national market into global supply chain market by considering connection among global markets. Canel and Das [4] developed an integrative mathematical model to connect global manufacturing and marketing. Jiang and Suer [11] compared two supply chain strategies in their work. The paper is based on their work and hybrid supply chain strategy is developed as an alternative approach. Ates [1] designed supply chain for the same product category and later simulated competition among four global manufacturers.

Group Technology (GT) was introduced to improve productivity in the Cellular Manufacturing System (CMS). Hyer and Wemmerlov [10] discussed Group

Technology in cellular manufacturing. Rajamani, Singh and Aneja [17] also mentioned that GT played an important role in cellular manufacturing. Singh and Mohanty [18] observed that not many works in the literature used fuzzy concepts to deal with multi-objective framework in the process planning. Chen and Cheng [5] proposed a supplementary procedure to solve the limitation of Adaptive Resonance Theory (ART). They mentioned that the performance of ART depended on the initial matrix of bottleneck process. The proposed supplementary procedures could improve the reliability of results. Moreover, a new mathematical model based on cell utilization was conducted by Mahdavi, Javadi, Fallah-Alipour and Slomp [15]. A comparison of part-machine grouping from this proposed method with the mathematical model from Chen and Cheng [5] was tested. The model from Mahdavi, Javadi, Fallah-Alipour and Slomp [15] tended to produce better results. A mixed integer non-linear model was analyzed by Bulgak and Bektas [3] for CMS. In this paper, the proposed model was an integrated approach to combine production planning and system reconfiguration. This CMS model was a new model, which includes sequence, duplicate machines, capacity of machines and lot splitting.

The literature reviews discussed so far included the deterministic CMS problem. However, cellular manufacturing is difficult to design in the real world due to uncertainty of manufacturing process. In order to deal with the uncertainty of product demand along with processing time, another research is proposed by Süer, Huang and Maddisetty [19]. A heuristic methodology was conducted to distinguish cell types in the CMS - Dedicated Cell (DC), Shared Cell (SC) and Remainder Cell (RC). The product family configuration and cell allocation are accomplished by mathematical analysis. The designed manufacturing system turned to successfully solve the uncertainty of product demand and processing time through simulation method. The methodology conducted by Süer, Huang and Maddisetty [19] is implemented in the current research for the purpose of designing the manufacturing system given the market demand, part-family formations, and the operations required to process the products. Erenay, Suer, Jing [8] also developed a mathematical model for designing alternative layered cellular system and they compared both works.

3 Problem Definition

In this research, a blood glucose test strip manufacturer is studied and alternative global supply chain strategies are compared. The strategies compared are, (1) independent facilities per region (host market strategy) and (2) newly proposed hybrid approach that allows family-based cell sharing across multiple facilities. Customers from three continents are considered in this study and they are Europe, Asia and North America. Three manufacturing facilities are assumed to produce these products and they are located in Ireland, China and Puerto Rico. The production data and manufacturing processes are obtained from Lobo [12]. All of the data are converted into common units by considering market share, revenue, and product price from Ates [1].

In most manufacturing systems, different products require to be processed on different machines. Due to high product variety, products are grouped into several families based on their similarity. Table 1 shows an example of product-machine incidence matrix. In this table, “1” in row i and column j indicates that product i needs to be produced on machine j. For example, Product 2 ($P2$) is processed on Machine 1 ($M1$) and Machine 3 ($M3$). One can observe that products with similar manufacturing processes are grouped together. Table 2 shows families and cells they are assigned to for this example.

However, in real life manufacturing systems, some product families may have quite high demand, which means they cannot be produced in one cell. Table 3 shows this multiple cell production system. For example, due to high demand, product families 1, 2 and 3 may need 2, 2 and 3 cells, respectively.

In many cases, demand values for product families follow a probabilistic distribution. In these situations, expected utilization for some cells of families may be low and as a result it may be desirable to share one cell by multiple families. A Dedicated Cell (DC) is assigned to one product family. A Shared Cell (SC) can run two product families, which have relatively similar operations. A Remainder Cell (RC) handles more than two product families. Both Shared Cells and Remainder Cells usually handle product families that have medium or low expected utilization values for some of its cells. Table 3 shows the cell sharing between three product families. For example, Cell 1 ($C1$) is Dedicated Cell for Product Family 1 ($F1$). C2 is also Dedicated Cell for $F2$. $C4$ is a Remainder Cell to be shared by $F1$, $F2$ and $F3$. Finally, $C3$ is a Shared Cell between $F1$ and $F2$ (Table 4).

Table 1. An example of product-machine incidence matrix

Product	Machines					
	$M1$	$M2$	$M3$	$M4$	$M5$	$M6$
$P1$	1	1	1	-	-	-
$P2$	1	-	1	-	-	-
$P3$	1	1	1	-	-	-
$P4$	-	-	-	1	1	-
$P5$	-	-	-	1	1	-
$P6$	1	-	1	-	-	1
$P7$	1	-	1	-	-	1

Table 2. Product families and cells

Family	Products	Cell	Machines in the Cell
$F1$	$P1$, $P2$, $P3$	Cell1	$M1$, $M2$, $M3$
$F2$	$P4$, $P5$	Cell2	$M4$, $M5$
$F3$	$P6$, $P7$	Cell3	$M1$, $M3$, $M6$

Table 3. Family vs. multiple cells due to high demand

Family	Cells						
	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$
$F1$	1	1	-	-	-	-	-
$F2$	-	-	1	1	-	-	-
$F3$	-	-	-	-	1	1	1

Table 4. Layered cellular design due to stochastic demand

Family	Cells			
	$C1$	$C2$	$C3$	$C4$
$F1$	1	-	1	1
$F2$	-	1	1	1
$F3$	-	-	-	1
	(DC)	(DC)	(SC)	(RC)

4 Alternative Supply Chain Designs

In this section, two alternative supply chain design strategies are discussed. Strategy 1 is the host market strategy where each facility in a continent covers the demand of the same continent. Strategy 2 discusses the proposed hybrid global supply chain strategy where cell sharing opportunities are explored across various facilities with product families in mind.

4.1 Strategy 1: Independent Supply Chain Design

In this strategy, each region produces many types of products to meet the demand of its own demand. Products are produced independently in different facilities, which lead to no transportation and information sharing between different regions. Figure 3 shows that the blood sugar strips are produced in three manufacturing facilities-China, Ireland and Puerto Rico [11].

4.2 Strategy 2: Hybrid Global Supply Chain Design

In this strategy, each region produces some of their own products. For example, products in Family 1 and Family 2 are produced in their own facilities in their continents. However, products in family 3 are produced in a single location and distributed to the remaining regions. Obviously, manufacturing cell(s) producing products in $F3$ are shared by various regions while this requires transportation of some products (Fig. 4).

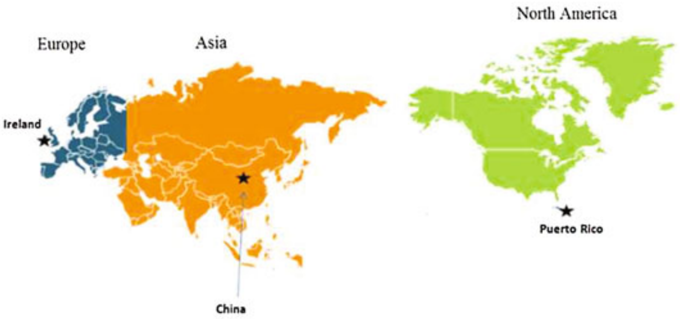


Fig. 3. Independent manufacturing systems

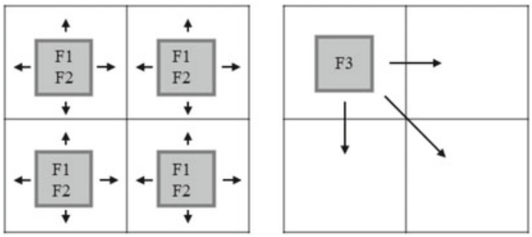


Fig. 4. The proposed hybrid global supply strategy

5 The Methodology Used

In each manufacturing facility, there are both fabrication and packaging cells. Products have been already divided into five families and this is out of scope of the paper. Two alternative strategies are discussed. Most of the computations are identical for both strategies. The only difference is how shared and remainder cells are created. To be more specific, in the host market strategy, DC, SC and RC are defined for each facility independently while, in the newly proposed strategy these opportunities are sought across multiple facilities potentially located in different continents. The general methodology is illustrated in Fig. 5. Cell utilizations are calculated by using cell capacity, product demand, etc. Each cell capacity is assumed 2000 h annually.

6 Mean Capacity Requirements and Standard Deviation

Historical demand values of four companies - Roche, LifeScan, Bayer and Abbott from 2002 to 2010 are used to forecast the 2011 demand [1]. In this research, demand values for families are assumed to follow normally distribution. Standard Deviation (σ) values for each product family are generated as a percentage of the corresponding mean demand (20%–25%). From [1], mean demand by family in all markets is shown in Table 5 along with Standard Deviations (σ).

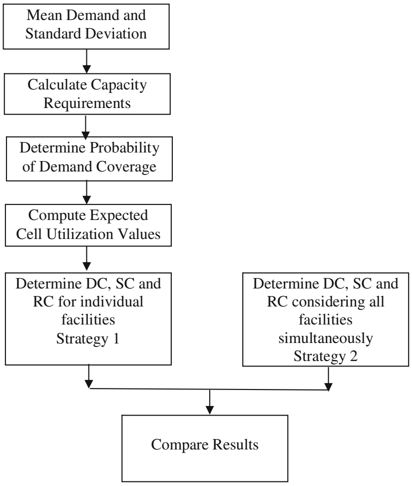


Fig. 5. The general methodology followed in this paper

Table 5. Mean demand and standard deviation by family in three regions

Family	China (C)			Ireland (I)			Puerto Rico (PR)		
	Mean	%	STDEV	Mean	%	STDEV	Mean	%	STDEV
1	1,422,286	25	355,571	1,422,925	25	355,731	1,337,087	25	334,271
2	7,098,188	24	1,703,565	7,101,377	24	1,704,330	6,672,986	24	1,601,516
3	6,711,423	24	1,610,741	6,714,438	24	1,611,465	6,309,389	24	1,514,253
4	24,313,261	21	5,105,784	24,324,183	21	5,108,078	22,856,826	21	4,799,933
5	3,137,454	25	784,363	3,138,863	25	784,715	2,949,511	25	737,377

Once mean demand and standard deviation values are known, the mean capacity requirements by product family are calculated by using Eq. (1) in [14]. Bottleneck Processing Time is defined as the longest processing time among operations in the cell.

$$\mathrm{MCR}_{\mathrm{FamilyNo}} = \mathrm{Mean}_{\mathrm{Demand}} \times \frac{\text{Bottleneck Processing Time}}{60}(hr). \tag{1}$$

For example, the mean capacity requirements for Product Family 2 in the manufacturing system of China region is decided by Mean Demand of Product Family 2 in China region which is 7,098,188. BPT (Bottleneck Processing Time) is $1/80 = 0.0125$ min in China region. The results of Mean Capacity Requirements and Standard deviation for different regions are shown in Table 6.

$$\mathrm{MCR}_{F2} = 7,098,188 \times \frac{0.0125}{60} = 1479 \text{ h},$$

$$\text{Standard Deviation}_{\mathrm{Capacity}_{F2}} = \sqrt{1,703,565^2 \times \frac{0.0125^2}{3600}} = 335.$$

Table 6. Mean demand and standard deviation by family in three regions

Family	China (C)		Ireland (I)		Puerto Rico (PR)	
	MCR	STDEV	MCR	STDEV	MCR	$\text{STDEV}_{\text{pity}}$
1	296	74	296	74	279	70
2	1479	355	1479	355	1390	334
3	1868	448	1869	449	1756	421
4	5065	1064	5068	1064	4762	1000
5	654	163	654	163	614	154

7 Calculating Demand Coverage Probabilities

The demand coverage probability is calculated for various cell levels. It shows the probability that a given number of cells will meet the demand. Capacity Requirements are assumed to follow the normal distribution. Each cell is available 2000 h per year. Demand Coverage Probability (DCP) for a family and cell combination is calculated by Eq. (2).

$$\text{DCP}_{\text{FamilyCell}} = \text{Normsdist}\left(\frac{2000 \times \text{CellNo.} - \text{MCR}_{\text{FamilyNo}}}{\text{STDEV}_{\text{Capacity}}}\right). \tag{2}$$

Demand Coverage Probability for all cells and families are calculated based on Mean Capacity Requirement and Standard Deviation. The Mean Capacity Requirement for Product Family 2 for China region is 1479 and Standard Deviation is 355. Based on these values, the Demand Coverage Probability for the first cell for family 2 is 93%. In other words, one cell covers demand at a very high level.

$$\text{DCP}_{F1C1} = \text{Normsdist}\left(\frac{2000 \times 1 - 1479}{335}\right) = 0.93.$$

All of the results of Demand Coverage Probability for different regions are shown in Table 7. For family 2 in China facility, by adding a second cell, the Demand Coverage Probability jumps to 99.99%.

7.1 Expected Cell Utilization Calculation

Expected Cell Utilization is determined by using Demand Coverage Probability, Mean Capacity Requirements and Standard Deviation as shown in Eq. (3)

$$\begin{aligned} E(C = X) = & P(CR > X) \times PU_1 + PX - 1 \leq CR \leq X \times PU_2 \\ & + P(CR < X - 1)PU_3, \end{aligned} \tag{3}$$

Table 7. Demand coverage probabilities

Family	China (C)					Ireland (I)					Puerto Rico (PR)				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	1.00	-	-	-	-	1.00	-	-	-	-	1.00	-	-	-	-
2	0.93	1.00	-	-	-	0.93	1.00	-	-	-	0.97	1.00	-	-	-
3	0.62	1.00	-	-	-	0.61	1.00	-	-	-	0.72	1.00	-	-	-
4	0.00	0.16	0.81	0.99	1.00	0.00	0.16	0.81	0.99	1.00	0.00	0.22	0.89	0.99	1.00
5	1.00	-	-	-	-	1.00	-	-	-	-	1.00	-	-	-	-

where

$E(C = X)$: Expected cell utilization for the X^{th} cell in a product family;
$P(CR > X)$: Probability that Capacity Requirements (CR)exceed $> X$;
PU_1: Percentage utilization of the X^{th} cell when $CR > X$, $PU_1 = 1.0$;
$P(X-1 \leq CR \leq X)$: Probability that CR between $X - 1$ and X;
PU_2: Percentage utilization of X^{th} cell when CR between $X-1$ and X;
$P(CR < X-1)$: Probability that $CR < X - 1$;
PPU_3: Percentage utilization of X^{th} cell when $CR < X - 1$, $PU_3 = 0.0$

PU_2 is solved by Eq. (4).

$$PU_2 = \int_{2000(X-1)}^{2000X} \frac{y \times f(y)}{2000 \times A} dy - (X - 1), \tag{4}$$

where

y: Variable represents CR;
$f(y)$: Probability density formation for CR;
A: Probability that CR between $X - 1$ and X

$f(y)$ and A are calculated by Eqs. (5) and (6)

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2 \frac{1}{2\sigma^2}}, \tag{5}$$

$$A = P(X - 1 \leq CR \leq X). \tag{6}$$

For example, Expected Cell Utilization of Product Family 1 for China region is decided by Probability that the Cell Required number larger than 1, Percentage utilization of 1st cell when $CR > 1$, Probability that CR between 0 and 1 and Percentage utilization of 1st cell when CR between 0 and 1.

$$
\begin{aligned}
E(C=1) &= P(CR > 2000) \times PU_1 + P(0 \le CR \le 2000) \times PU_2 + P(CR < 0) \times PU_3 \\
&= 0 \times 1 + 1 \times PU_2, \\
PU_2 &= \int_0^{2000} \frac{y}{2000 \times 1} \times \frac{1}{74 \times \sqrt{2\pi}} e^{-(y-296)^2 \times \frac{1}{2 \times 74^2} dy} = 0.1480.
\end{aligned}
$$

All of the results of Expected Cell Utilization calculations for different regions are given in Table 8.

8 The Heuristic Algorithm for Layered Cellular Design

Having determined Expected Cell Utilization values for all possible cell segments, Dedicated Cells (DC), Shared Cells (SC), and Remainder Cells (RC) are identified using the heuristic algorithm described briefly here. The heuristic algorithm by Süer, Huang and Maddisetty [19] is used for identifying cell types. Expected Cell Utilization values are sorted in the decreasing order first. The assignment starts with the highest Expected Cell Utilization. If the expected cell utilization is 100%, this cell is considered to be a Dedicated Cell (DC). If the expected cell utilization is larger than 50%, a cell will be allocated to a product family. Then other similar product families are allocated to the cell to make the cell coverage close to 100% by considering similarities among families. These cells are named Shared Cells (SC) if they process only two product families and remainder cells if they process three or more product families.

Threshold value is the lowest acceptable similarity coefficient that allows two families to be grouped in a cell. The Similarity Threshold is set to 77%

Table 8. Demand coverage probabilities

Family	China (C)					Ireland (I)					Puerto Rico (PR)				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	0.15	0.00	-	-	-	0.15	0.00	-	-	-	0.14	0.00	-	-	-
2	0.73	0.01	0.00	-	-	0.73	0.01	0.00	-	-	0.69	0.00	-	-	-
3	0.87	0.06	0.00	-	-	0.87	0.06	0.00	-	-	0.84	0.04	0.00	-	-
4	1.00	0.96	0.52	0.06	0.00	1.00	0.96	0.52	0.06	0.00	1.00	0.94	0.42	0.03	0.00
5	0.33	0.00	-	-	-	0.33	0.00	-	-	-	0.31	0.00	-	-	-

Table 9. Similarity coefficients between product families

Family	1	2	3	4	5
1	-	1.00	0.89	0.78	0.70
2	1.00	-	0.89	0.78	0.70
3	0.89	0.89	-	0.70	0.80
4	0.78	0.78	0.70	-	0.89
5	0.70	0.70	0.80	0.89	-

Table 10. Cell Types for independent regions

Family	China (C)					Ireland (I)					Puerto Rico (PR)				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	-	-	-	0.15	-	-	-	-	0.15	-	-	-	-	0.14	-
2	-	0.01	-	0.73	-	-	0.01	-	0.73	-	-	-	-	0.69	-
3	-	-	0.87	0.06	-	-		0.87	0.06	-	-	-	0.84	0.04	-
4	1.00	0.96	-	0.06	0.52	1.00	0.92	-	0.02	0.52	1.00	0.94	-	0.03	0.42
5	-	-	-	-	0.33	-	-	-	-	0.33	-	-	-	-	0.31
	DC	SC	DC	RC	SC	DC	SC	DC	RC	SC	DC	DC	DC	RC	SC

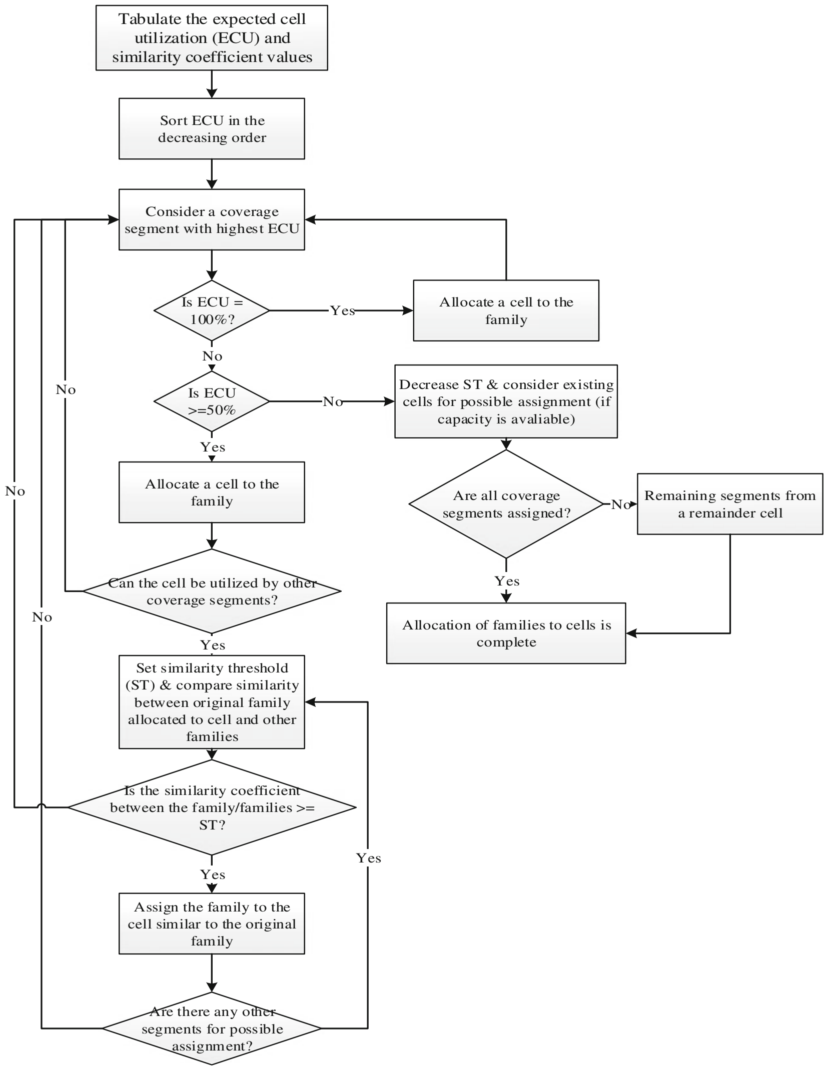


Fig. 6. The heuristic procedure for the layered cellular design

Table 11. Cell types for China based on new approach

Family	China					
	1	2	3	4	5(C + PR)	6(C + I + PR)
1	-	-	-	0.15	-	-
2	-	0.01	-	0.73	-	-
3	-	-	0.87	0.06	-	-
4	1.00	0.96	-	0.06	0.52 + 0.42	-
5	-	-	-	-	-	0.33 + 0.33 + 0.31
	DC	SC	DC	RC	DC	DC

Table 12. Cell types for Puerto Rico and Ireland based on new approach

Family	Puerto Rico (PR)				Ireland (I)				
	1	2	3	4	1	2	3	4	5
1	-	-	-	0.14	-	-	-	0.15	-
2	-	-	-	0.69	-	0.01	-	0.73	-
3	-	-	0.84	0.04	-	-	0.87	0.06	-
4	1.00	0.94	-	0.03	1.00	0.92	-	0.02	0.52
5	-	-	-	-	-	-	-	-	-
	DC	DC	DC	DC DC	DC	DC	DC	DC	

in this paper. For example, when sorting the *ECU* for Ireland region, the highest *ECU* is 100% for Family 4 in Cell 1. Then Product family 4 is allocated to cell 1. When the second cell with 92% utilization is considered, it is allocated to a new $cell - Cell2$. Table 10 is used to search the similar Families with Product Family 4. From Table 9, Families 1, 2 and 5 can be considered to share a cell with Product Family 4 (similarity coefficient > 0.77). However, merging this cell with Family 1 or Family 5 will exceed 10% utilization. The only option is to merge Cell 2 (1% utilization) of Family 2 with Family 4. In Ireland case, there are two Dedicated Cells, two Shared Cells and one Remainder Cell. The results are summarized in Table 10. The heuristic procedure is given in Fig. 6.

9 The Proposed Hybrid Global Supply Approach

The proposed approach looks into global supply chain and explores the opportunities for forming shared cells and remainder cells across global facilities. In this paper, a new cell (cell 6) is opened in China where the entire Family 5 is produced to meet the global demand. In the same facility, demand for product family 4 original assigned to Puerto Rico facility is re-assigned to China facility. The Puerto Rico facility is thus reduced to four cells with three DCs and one RC. Ireland facility is also changed to 3 DCs, one SC and one RC as shown in

Table 13. Comparison of results

Parameters	Independent supply chain				New hybrid supply chain
	China	Ireland	P.R.	Total	
# of DCs	2	2	2	6	9
# of SCs	2	2	2	6	3
# of RCs	1	1	1	3	3
# of total cells	5	5	5	15	15
# of Cell Util > 0.9	3	3	3	9	11
# of workers	82	82	82	246	234
# of machines	66	66	66	198	198

Table 12. In short, China facility becomes a global supplier for Product Family 5 and partial global supplier for Product Family 4. All three facilities supply their own markets with respect to product Families 1, 2 and 3 (Table 11).

10 Conclusion

The comparison between two designs is presented in Table 13. It shows that total number remains the same for both designs. However, the number of workers is reduced in the new design. This is due to balancing issues in the cells. In other words, each cell is equipped with all the machines needed to run all the families (maximum number of machines) while some of them remained idle while running different families. Another benefit was that there were more dedicated cells under the new design. Furthermore, number of cells with a utilization value greater than 0.90 was higher. It is expected that flowtimes will be shorter in dedicated cells. This in turn reduces the work-in-process inventory. For a more complete analysis, transportation costs, labor costs as well as investments costs have to be included. These results could not be included due to space concerns.

References

1. Ates O (2013) Tglobal supply chain and competitive business strategies: a case study of blood sugar monitoring industry. https://etd.ohiolink.edu/
2. Brush TH, Marutan CA, Karnani A (1999) The plant location decision in multinational manufacturing firms: an empirical analysis of international business and manufacturing strategy perspectives. Prod Oper Manag 8:109–132
3. Bulgak AA, Bektas T (2009) Integrated cellular manufacturing systems design with production planning and dynamic system reconfiguration. Eur J Oper Res 192:414–428
4. Canel C, Das SR (2002) Modeling global facility location decisions: integrating marketing and manufacturing decisions. Ind Manag Data Syst 102:110–118
5. Chen SJ, Cheng CS (1995) A neural network-based cell formation algorithm in cellular manufacturing. Int J Prod Res 33:20–542

6. Chopra S, Meindl P (2013) Supply chain management: strategy, planning & operations, 2nd edn. Prentice Hall, Upper Saddle River
7. Dicken P (1992) Global shift: the internationalization of economic activity, 2nd edn. Paul Chapman Publishing, London
8. Erenay B, Suer GA et al (2015) Comparison of layered cellular manufacturing system design approaches. Comput Ind Eng 85:346–358
9. Haug P (1992) An international location and production transfer model for high technology multinational enterprises. Int J Prod Res 30:559–572
10. Hyer NL (1992) Education: Mrp/gt: a framework for production planning and control of cellular manufacturing. Decis Sci 13:681–701
11. Jiang J, Suer GA (2016) Alternative global supply chain design strategies for a blood sugar strip manufacturer considering layered cellular design. In: Proceedings of the global joint conference on industrial engineering and its application areas 2016, Istanbul, Turkey, 14–15 July, pp 14–15
12. Lobo R (2006) Comparison of connected vs disconnected cellular systems using simulation. Ph.D. thesis, Ohio University
13. Lowe TJ, Wendell RE, Hu G (2002) Screening location strategies to reduce exchange rate risk. Eur J Oper Res 136(3):573–590
14. Maddisetty S (2005) Design of shared cells in a probabilistic demand environment. Ph.D. thesis, Ohio University
15. Mahdavi I, Javadi B et al (2007) Designing a new mathematical model for cellular manufacturing system based on cell utilization. Appl Math Comput 190:662–670
16. New SJ, Payne P (1995) Research frameworks in logistics: three models, seven dinners and a survey. Int J Phys Distrib Logistics Manag 25:60–77
17. Rajamani D, Singh N, Aneja YP (1990) Integrated design of cellular manufacturing systems in the presence of alternative process plans. Int J Prod Res 28:1541–1554 (in Chinese)
18. Singh N, Mohanty BK (1991) A fuzzy approach to multi-objective routing problem with applications to process planning in manufacturing systems. Int J Prod Res 29:1161–1170
19. Süer GA, Huang J, Maddisetty S (2010) Design of dedicated, shared and remainder cells in a probabilistic demand environment. Int J Prod Res 48:92–114

Design of Closed-Loop Supply Chain Model with Efficient Operation Strategy

Xing Chen[1], Anudari Chuluunsukh[1], YoungSu Yun[1(✉)], and Mitsuo Gen[2]

[1] Department of Business Administration, Chosun University, Gwangju, Korea
ysyun@chosun.ac.kr
[2] Fuzzy Logic Systems Institute, Tokyo University of Science, Katsushika, Japan

Abstract. In this paper, a closed-loop supply chain model with efficient operation strategy (CLSC-OS) is proposed. In the CLSC-OS, various facilities which can be used in forward logistics (FL) and reverse logistics (RL) are taken into consideration. A mathematical formulation is proposed to design the CLSC-OS and it is implemented using genetic algorithm (GA) approach. In numerical experiment, for efficient operation strategy, several scenarios using various factors such as remanufacturing rate, profit premium rate, discount rate, and return rate based on revenue-cost rate is considered and analyzed. Experimental results shows that discount rate and profit premium rate have a significant influence on revenue-cost rate.

Keywords: Closed-loop supply chain · Efficient operation strategy · Forward logistics · Reverse logistics · Genetic algorithm

1 Introduction

In general, designing closed loop supply chain (CLSC) model is to maximize the utilization of resources. Especially, return management such as reuse and resale of the recovered parts and products has been focused on many companies. For example, by 1997, approximately 73,000 U.S. companies had earned $53 billion though resale activities of remanufactured products [1]. This means that return management using used product becomes more and more important.

Return management using used product is usually consisted of three aspects. The first aspect is the operational management of returned product, the second one is that of remanufacturing activity, and the third one is that of secondary market [6]. Ozkir and Basligil considered the operational management of returned product [8]. They found that regulating the rate of returned product can increase the total profit and revenue. Subramoniam found that remanufacturing activities can increase profit, save resource and energy, create new market [9]. Van Daniel et al. suggested that con-sidering both new market with new product activity and secondary market with resale activity can get more profit than the new market alone [7].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_79

However, unfortunately, little attention to the relationship among the three aspects of return management has been done in conventional studies. Therefore, in this paper, we consider the relationship based on various scenarios using remanufacturing rate, profit premium rate, discount rate, and return rate and analyze which factor can be an important one to increase total profit.

For achieving this objective, a CLSC model with efficient operation strategy (CLSC-OS) is proposed in Sect. 2. The CLSC-OS is consisted of various facilities such as supplier, product manufacturer, retailer, recovery center, and secondary market at each stage. In Sect. 3, a mathematical formulation is proposed for implementing the CLSC-OS. The mathematical formulation is represented as a nonlinear mixed integer programming (NIMIP). The NIMIP is implemented using a genetic algorithm (GA) approach in Sect. 4. In numerical experiment of Sect. 5, the CLSC-OS is presented and various scenarios using the CLSC-OS is considered. For various scenarios, remanufacturing rate, profit premium rate, discount rate, and return rate are considered and they all are used for analyzing the revenue-cost rate resulting from the implementation of the CLSC-OS. Some conclusions and a room for improvement to future study are followed in Sect. 6.

2 Proposed CLSC-OS

The proposed CLSC-OS is an integrated model combining forward logistics (FL) and reverse logistics (RL), which is consisted of suppliers in areas 1, 2, 3, and 4, product manufacturer, part inventory distribution center, product distribution center and retailer in FL and customer, collection center, recovery center, secondary market and waste disposal center in RL. For effective use of the recovered parts from recovery center, part inventory distribution center is also taken into consideration. The conceptual network structure of the proposed CLSC-OS is shown in Fig. 1.

In Fig. 1, the production and recovery flows are as follows. New part types 1, 2, 3, and 4 (NP1, NP2, NP3, and NP4) are produced at the suppliers of areas 1, 2, 3, and 4, respectively. NP1, NP2, NP3, and NP4 are then sent to part inventory distribution center. Recovered parts (RP1, RP2, RP3, and RP4) from recovery center are also sent to part inventory distribution center. Product manufacturer produces product using NP1, NP2, NP3, NP4, RP1, RP2, RP3, and RP4 from part inventory distribution center. The product is sent to retailer through product distribution center and sold to customer. The returned product from customer is collected at collection center and then sent to recovery center. At recovery center, the returned product is checked and classified into recoverable and unrecoverable products. The quality of the recoverable product with a_1% is recovered at recovery center and then resold at secondary market. The unrecoverable product is disassembled into recoverable and unrecoverable parts. The quality of the recoverable part with a_2% is recovered at recovery center and then sent to part inventory distribution center. The unrecoverable part with a_3% is sent to waste disposal center to be burned or landfilled.

Especially, in the proposed CLSC-OS, the part inventory distribution center can regulate the transportation amount of NP1, NP2, NP3, NP4, RP1, RP2,

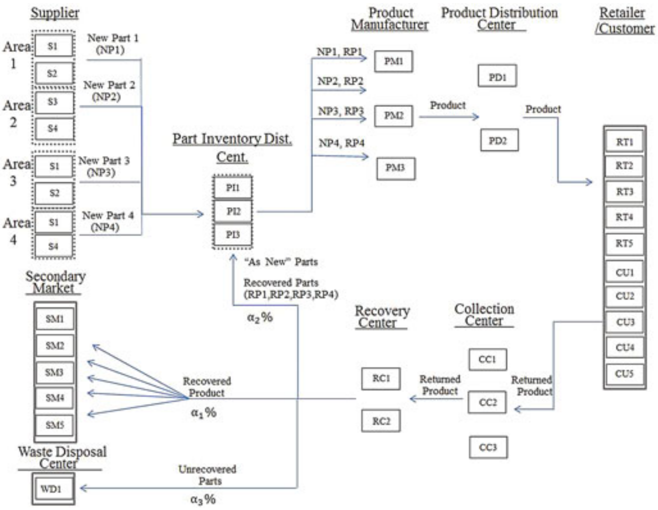


Fig. 1. The conceptual network structure of the CLSC-OS

RP3, and RP4. For instance, if product manufacturer want to produce 100 products and a_2 is 10% (10 recovered parts = 100 returned product $\times$10%), then part inventory distribution center sends 90 NP1, 90 NP2, 90 NP3, and 90 NP4 as new parts and 10 RP1, 10 RP2, 10 RP3, and RP4 as recovered parts to product manufacturer.

3 Mathematical Formulation

Beofre suggesting the mathematical formulation for the CLSC-OS, the following assumptions should be considered.

- Single product is produced within a fixed period.
- New market (retailer) and resale market (secondary market) are separated.
- The qualities of the new parts (NP1, NP2, NP3, and NP4) from supplier and those of the recovered parts (RP1, RP2, RP3, and RP4) from recovery center are identical.
- All products (new and recovered products) at retailer and secondary market are sold to customer.
- The number of facilities at each stage is known.
- All facilities at retail/customer, secondary market, and waste disposal center should be opened. However, only one facility at supplier, product manufacturer, product distribution center, part inventory distribution center, collection center, and recovery center should be opened, which means that the other facilities at each stage should be closed.
- Fixed costs of all facilities are different and known.
- Unit handling costs of all facilities at same stage are identical and known.

- Unit transportation costs of all facilities at each stageare different and known.
- The handling capacity of the facilities at a stage is the same or greater than that of previous ones.
- The return rate from customer is fixed.
- The discount rates of recovered product and part are calculated according to the quality of returned product and part.

The index sets, parameters, and decision variables for the mathematical formulation are as follows.
Index set

a : index of area at supplier; $a \in A$;
b : index of new product; $b \in B$;
h : index of supplier; $h \in H$;
i : index of part inventory distribution center; $i \in I$;
j : index of product manufacturer; $j \in J$;
k : index of product distribution center; $k \in K$;
l : index of retailer/customer; $l \in L$;
m : index of collection center; $m \in M$;
n : index of recovery center; $n \in N$;
o : index of secondary market; $o \in O$;
p : index of waste disposal center; $p \in P$

Parameter

FDS_{ha} : fixed cost at supplier h of area a;
FDI_i : fixed cost at part inventory distribution center i;
FDO_j : fixed cost at product manufacturer j;
FDD_k : fixed cost at distribution center k;
FDC_m : fixed cost at collection center m;
FDR_n : fixed cost at recovery center n;
UHS_{ha} : unit handling cost at supplier h of area a;
UHO_j : unit handling cost at product manufacturer j;
UHD_k : unit handling cost at product distribution center k;
UHC_m : unit handling cost at collection center m;
UHR_n : unit handling cost at recovery center n;
UHE_e : unit handling cost at secondary market e;
UHW_p : unit handling cost at waste disposal center p;
UTS_{hai} : unit transportation cost from supplier h of area a to part inventory distribution center i;
UTI_{ij} : unit transportation cost from part inventory center i to product manufacturer j;
UTO_{jk} : unit transportation cost from product manufacturer j to product distribution center k;
UTD_{kl} : unit transportation cost from product distribution center k to retailer/customer l;

UTA_{lm} : unit transportation cost from retailer/customer l to collection center m;
UTC_{mn} : unit transportation cost from collection center m to recovery center n;
UTE_{no} : unit transportation cost from recovery center n to secondary market o;
UTI_{ni} : unit transportation cost from recovery center n to part inventory distribution center i;
UTL_{np} : unit transportation cost from recovery center n to waste disposal center p;
KCI_{i} : holding cost at part inventory distribution center i;
UPN_{b} : unit price of new product b;
PRN_{b} : profit premium rate for new product b at retailer l

Decision variables
TCs_{ha} : treatment capacity at supplier h of area a
TCi_{i} : treatment capacity at part inventory distribution center i
TCo_{j} : treatment capacity at product manufacturer j
TCd_{k} : treatment capacity at product distribution center k
TCu_{l} : treatment capacity at retailer/customer l
TCc_{m} : treatment capacity at collection center m
TCe_{n} : treatment capacity for recovered product at recovery center n
TCi_{n} : treatment capacity for recovered part at recovery center n
TCl_{n} : treatment capacity for unrecovered part at recovery center n
TCe_{e} : treatment capacity at secondary market e
TCw_{p} : treatment capacity at waste disposal center p
Q : quantity of new product
Rt : return rate
DS_o : discount rate for recovered product
DS_i : discount rate for recovered part
ys_{ha} : takes the value of 1 if suppler h at area a is opened and 0 otherwise
yi_i : takes the value of 1 if part inventory distribution center i is opened and 0 otherwise
yo_j : takes the value of 1 if product manufacturer j is opened and 0 otherwise
yd_k : takes the value of 1 if product distribution center k is opened and 0 otherwise
yc_m : takes the value of 1 if collection center m is opened and 0 otherwise
yr_n : takes the value of 1 if recovery center n is opened and 0 otherwise

Using the above mentioned index sets, parameters, and decision variables, the mathematical formulation is suggested as follows:

Maximize Total Revenue (TR)

$$\begin{aligned}\mathrm{TR} = {} & \text{revenue of new product} + \text{revenue of recovered product} \\ & + \text{revenue of recovered part}\end{aligned} \tag{1}$$

$$= \sum_i Q \times \mathrm{PRN}_b + \sum_o Q \times \mathrm{UPN}_b \times \alpha_1 \times DS_o + \sum_o Q \times \mathrm{UPN}_b \times \alpha_2 \times DS_i, \quad (2)$$

where

$$\mathrm{UPN}_b = (\mathrm{TCF} \times (1 + \mathrm{PRN}_b))/Q. \quad (3)$$

Minimize Total Cost (TC)

$$\mathrm{TC} = \text{total cost of FL (TCF)} + \text{total cost of RL (TCR)}. \quad (4)$$

$$\begin{aligned}
\mathrm{TCF} &= \text{total FL fixed cost} + \text{ total FL handling cost} \\
&\quad + \text{total FL transportation cost} \\
&= \sum_a \sum_h (\mathrm{FDS}_{ha} \times ys_{ha}) + \sum_i (\mathrm{FDI}_i \times yi_i) + \sum_i (\mathrm{FDO}_j \times yo_j) \\
&\quad + \sum_a \sum_h (\mathrm{UHS}_{ha} \times TCs_{ha} \times ys_{ha}) + \sum_j (\mathrm{UHO}_j \times TCo_j \times yo_j) \\
&\quad + \sum_j (\mathrm{UHD}_k \times TCd_k \times yd_k) + \sum_a \sum_h \sum_i (UTS_{hai} \times TCs_{ha} \times ys_{ha} \times yi_i) \\
&\quad + \sum_{ij} (\mathrm{UTI}_{ij} \times TCi_i \times yi_i \times yo_j) + \sum_{jk} (\mathrm{UTO}_{jk} \times TCo_j \times yo_j \times yd_k) \\
&\quad + \sum_{kl} (\mathrm{UTD}_{kl} \times TCd_k \times yd_k),
\end{aligned} \quad (5)$$

$$\begin{aligned}
\mathrm{TCR} &= \text{total RL fixed cost} + \text{total RL handling cost} \\
&\quad + \text{total RL transportation cost} + \text{holding cost} \\
&= \sum_m (\mathrm{FDC}_m \times yc_m) + \sum_i (\mathrm{FDR}_n \times yr_n) + \sum_a \sum_h (\mathrm{UHC}_m \times TCc_m \times yc_m) \\
&\quad + \sum_n (\mathrm{UHR}_n \times TCe_n \times yr_n) + \sum_n (UHR_n \times TCi_n \times yr_n) \\
&\quad + \sum_n (\mathrm{UHR}_n \times TCl_n \times yr_n) + \sum_e (\mathrm{UHE}_e \times TCe_e) + \sum_p (\mathrm{UHW}_p \times TCw_p) \\
&\quad + \sum_{lm} (\mathrm{UTA}_{lm} \times TCu_l \times yc_m) + \sum_{mn} (\mathrm{UTC}_{mn} \times TCc_m \times yc_m \times yr_n) \\
&\quad + \sum_{no} (\mathrm{UTE}_{no} \times TCe_n \times yr_n) + \sum_{ni} (\mathrm{UTI}_{ni} \times TCi_n \times yr_n) \\
&\quad + \sum_{ni} (\mathrm{UTL}_{np} \times TCl_n \times yr_n) + \sum_i (\mathrm{KCI}_i \times yi_i)
\end{aligned} \quad (6)$$

$$\text{s.t. } \sum_h ys_{ha} = 1, \forall a \in A \tag{7}$$

$$\sum_i yi_i = 1 \tag{8}$$

$$\sum_j yo_j = 1 \tag{9}$$

$$\sum_k yd_k = 1 \tag{10}$$

$$\sum_m yc_m = 1 \tag{11}$$

$$\sum_n yr_n = 1 \tag{12}$$

$$\sum_h (UHS_{hai} \times ys_{ha}) - \sum_i (TCi_i \times yi_i) = 0, \forall a \in A \tag{13}$$

$$\sum_i (TCi_i \times yi_i) - \sum_j (UHO_j \times yo_j) = 0 \tag{14}$$

$$\sum_i (\mathrm{UHO}_j \times yo_j) - \sum_k (\mathrm{UHD}_k \times yd_k) = 0 \tag{15}$$

$$\sum_k (\mathrm{UHD}_k \times yd_k) - \sum_l TCu_l = 0 \tag{16}$$

$$\sum_l TCu_l - \sum_m (\mathrm{UHC}_m \times yc_m) = 0 \tag{17}$$

$$\sum_i (TCi_i \times yi_i) - \alpha_2 \sum_n (\mathrm{UHR}_n \times yr_n) \geq 0 \tag{18}$$

$$\sum_n (HCrc_n \times x\_rc_n) - \alpha_3 \sum_p (UHW_p \times yw_p) \geq 0 \tag{19}$$

$$\sum_n (\mathrm{UHR}_n \times yr_n) - \sum_o (\mathrm{UHC}_m \times yc_m) = 0 \tag{20}$$

$$\sum_n (\mathrm{UHR}_n \times yr_n) - \alpha_1 \sum_e \mathrm{UHE}_e \geq 0 \tag{21}$$

$$ys_{ha} = \{0, 1\}, \forall h \in H, a \in A \tag{22}$$

$$yi_i = \{0, 1\}, \forall i \in I \tag{23}$$

$$yo_j = \{0, 1\}, \forall j \in J \tag{24}$$

$$yd_k = \{0, 1\}, \forall k \in K \tag{25}$$

$$yc_m = \{0, 1\}, \forall m \in M \tag{26}$$

$$yr_n = \{0, 1\}, \forall n \in N \tag{27}$$

$$\begin{aligned} &\mathrm{UHS}_{hai}, \mathrm{UHO}_j, \mathrm{UHD}_k, \mathrm{UHC}_m, \mathrm{UHR}_n, \mathrm{UHE}_e, \mathrm{UHW}_p \geq 0 \\ &\forall h \in H, a \in A, \forall i \in I, j \in J, \forall k \in K, \\ &\forall m \in M, \forall n \in N, \forall e \in E, \forall p \in P \end{aligned} \tag{28}$$

The objective function is divided into two parts. First part is to maximize total revenue and second one is to minimize total cost. Maximizing the total revenue is represented in Eq. (2) and minimizing the total cost is in Eq. (4). Each objective function should be maximized and minimized under satisfying all constraints which represented from Eqs. (7)–(28). Equations (7)–(12) mean that only one facility at each stage should be opened. Equations (13)–(21) show that the capacity of each opened facility is the same or greater than that of the previous one. Equations (22)–(27) means that each decision variables can take the value 0 or 1. Equation (27) refers to non-negativity.

4 Genetic Algorithm (GA) Approach

Genetic Algorithm (GA) approach was proposed by Goldberg [5] and has been used to solve many optimization problems [3,4,10]. GA approach starts with producing a random initial population and each individual in the population represents a potential solution. The next step is that each individual is selected by evaluating the measure of fitness and producing a next new population. All individuals in the new population are undergoing crossover and mutation operators. For crossover and mutation opera-tors, various methodologies such as one-point crossover operator (1X), two-point crossover operator (2X), and random mutation operator have been developed. After applying the crossover and mutation operators to each individual, an offspring is produced and the fitness is evaluated. These procedure is repeated until a pre-defined stop condition is satisfied or optimal solution already known is located. Finally, the optimal solution is produced.

The GA approach [2,3] for implementing the mathematical formulation in Sect. 3 is as Fig. 2:

```
procedure: GA approach
input: problem data, GA parameters
output: best solution
begin
    t = 0;
    generate initial population randomly P(t) satisfying constraints;
    evaluate P(t) using fitness test by decoding routine;
    while (not termination condition) do
        create N(t) from P(t) by two cut-point crossover routine;
        create N(t) from P(t) by random mutation routine;
        reproduce P(t + 1) from P(t) and N(t) by elitist selection rou-
tine;
        t = t + 1
    end
    output best solution
end;
```

Fig. 2. The GA approach

5 Numerical Experiments

In numerical experiment, various scenarios are considered for evaluating the rate of the total revenue in Eq. (1) and the total cost in Eq. (4). First, a scale for the proposed CLSC-OS is presented in Table 1. As shown in Table 1, numbers of the suppliers (S) at each area, part inventory distribution center (PI), product manufacturer (PM), product distribution center (PD), retailer/customer (RT), collection center (CC), recovery center (RC), secondary market (SM), waste disposal center (WD) are displayed.

Table 1. Scale of the CLSC-OS

No. of S Area 1, 2, 3, 4	No. of PI	No. of PM	No. of PD	No. of RT	No. of CC	No. of RC	No. of SM	No. of WD
4	2	4	2	10	4	2	10	2

Table 2 shows various cases of remanufacturing for returned product at recovery center. a_1, a_2, and a_3 refer to the rates of recovered product, recovered part, and unre-covered part, respectively. For instance, Case 1 means that (i) 60% of all returned products are recovered and sent to secondary market, (ii) 30% of them are recovered and sent to part inventory distribution center and (iii) 10% of them are unrecovered and sent to waste disposal center.

Table 2. Various cases of remanufacturing

Case	a_1	a_2	a_3
1	0.6	0.3	0.1
2	0.6	0.2	0.2
3	0.6	0.1	0.3
4	0.7	0.2	0.1
5	0.7	0.1	0.2
6	0.8	0.1	0.1

Table 3 shows various scenarios. Each scenario has four factors using the case of remanufacturing, return rate, discount rate and profit premium rate. By the each scenario, various relationships between total revenue and total cost are compared. For instance, Scenario 1 indicates that various cases in Table 2 and various profit premium rates ($P_m = 0.4, 0.3, 0.2, 0.1,$ and 0.0) under the fixed values at return rate and discount rate are considered to solve the two mathematical formulations in Eqs. (1) and (4). The computation results are used for comparing the relationships between total revenue and total cost.

Each scenario in Tables 2 and 3 under the scale of the proposed CLSC-OS is implemented using GA approach. The parameters used in GA approach is that population size is 20, crossover rate 0.5, and mutation rate 0.3. The GA approach is executed under the following computation environment: Matlab R2015 under

Table 3. Various scenarios

Scenario	Case	Return rate	Discount rate	Profit premium rate
1	1, 2, 3, 4, 5, 6	1.0	0.5	0.4, 0.3, 0.2, 0.1, 0.0
2	1	1.0	0.5	0.4, 0.3, 0.2, 0.1, 0.0
3	1	1.0	0.5, 0.4, 0.3, 0.2	0.4, 0.3, 0.2, 0.1, 0.0
4	1, 2, 3, 4, 5, 6	0.9	0.5	0.4, 0.3, 0.2, 0.1, 0.0

IBM compatible PC 3.40GHZ processor (Inter Core i7-3770 CPU), 8GB and Window 7.

Figure 3 shows the computation results of total revenue-cost rates in Scenario 1. Each curve with different colors shows the changes of the total revenue-cost rates under the changes of profit premium rate. For instance, the top blue-colored curve shows a fixed values of 1.4 in the total revenue-cost rates of all cases under the fixed values of 0.4 profit premium rate, 100% return rate and 0.5 discount rate. Similar situations are also shown in the other colored-curves. Therefore, the results shown in Fig. 3 indicates that the changes of remanufacturing rate has no significant influence on the increase of total profit.

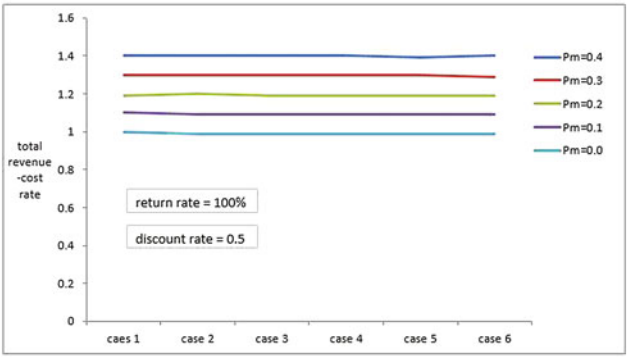


Fig. 3. Computation results of total revenue-cost rates in Scenario 1

Figure 4 shows the changes of the total revenue-cost rates under various profit premium rates, when return and discount rates are fixed at 100% and 0.5, respectively. As the changes from profit premium rate 0.4 to 0.0, the total revenue-cost rate is continuously decreasing. Therefore, the results shown in Fig. 4 indicate that the profit premium rate has significant influence on the increase of total profit.

Figure 5 shows the changes of the total revenue-cost rates for each profit premium rate under various discount rates, when return rate is fixed at 100%. For instance, the top blue-colored curve shows that total revenue-cost rates is

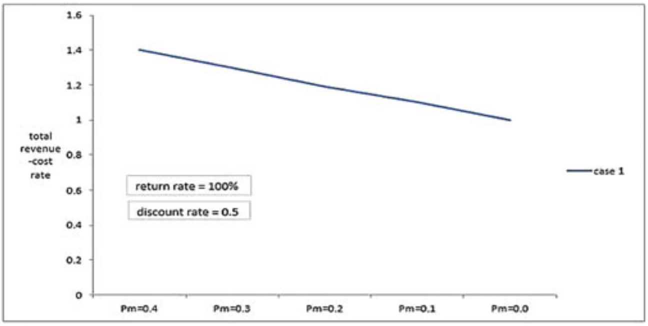


Fig. 4. Computation results of total revenue-cost rates in Scenario 2

increasing according to the changes of discount rate from 0.5 to 0.2. Similar situations are also shown in the other colored-curves. Therefore, the results shown in Fig. 5 indicate that the change of discount rates has significantly influence on the increase of total profit.

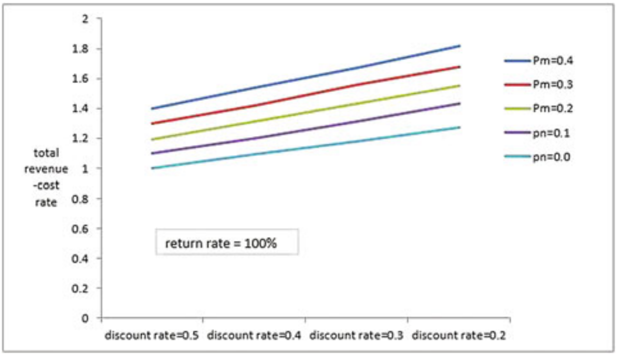


Fig. 5. Computation results of total revenue-cost rates in Scenario 3

Similar to Figs. 3 and 6 shows the computation results of total revenue-cost rates in Scenario 4. Each curve with different colors shows the changes of the total revenue-cost rates under the changes of profit premium rate. For instance, the top blue-colored curve shows a little change around 1.55 for each case when the discount rate and profit premium rate are fixed at 90% and 0.4, respectively. Similar curves are also shown in the profit premium rates 0.3 and 0.0. However, the curves of profit premium rates 0.2 and 0.1 do not show any changes for each case. Therefore, the results shown in Figs. 3 and 6 indicate that the changes of remanufacturing rate and return rates has no significant influence on the increase of total profit, though there is a little changes according to each case.

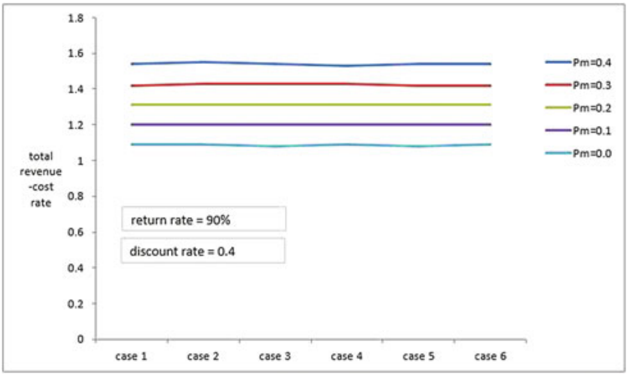


Fig. 6. Computation results of total revenue-cost rates in Scenario 4

6 Conclusion

The first objective of this study is to design a CLSC-OS with various facilities at each stage and the second one is to find a good alternative for increasing total profit under the changes of remanufacturing rate, discount rate, profit premium rate, and return rate.

For achieving the two objectives, the CLSC-OS has been proposed. The proposed CLSC-OS has supplies at areas 1, 2, 3, and 4, product manufacturer, product distribution center, and retailer in FL and customer, collection center, recovery center, secondary market and waste disposal center in RL. This is represented in a mathematical formulation which has two objectives of the maximization of total revenue and the minimization of total cost under various constraints. The mathematical formulation is implemented using GA approach. In numerical experiment, four types of scenarios with the changes of remanufacturing rate, discount rate, profit premium rate, and return rate have been considered and executed using GA approach. The experimental results have shown that the changes of the profit premium rates and the discount rates have no significant influence on the increase of total profit, but the changes of the profit premium rates and the discount rates have significant influence on the increase of total profit. Therefore, decreasing the profit premium rates and the discount rates can be a good alternative for increasing the total profit (= total revenue − total cost) in the proposed CLSC-OS.

However, in this paper, since only four scenarios have been considered, more various scenarios with larger-scaled CLSC-OS will be considered to compare the changes of remanufacturing rate, discount rate, profit premium rate, and return rate. This will be left to future study.

Acknowledgements. This work is partly supported by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS): No. 15K00357.

References

1. Atasu A, Sarvary M, Wassenhove LNV (2008) Remanufacturing as a marketing strategy. Manage Sci 54(10):1731–1746
2. Gen M, Cheng R (1997) Genetic algorithms and engineering design. Wiley, New York
3. Gen M, Cheng R (2000) Genetic algorithms and engineering optimization. Wiley, New York
4. Gen M, Zhang W et al (2016) Recent advances in hybrid evolutionary algorithms for multiobjective manufacturing scheduling. Comput Ind Eng doi:10.1016/j.cie.2016.12.045
5. Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning, vol 13. Addsion-Wesley, Reading
6. Guide VDR, Wassenhove LNV (2009) The evolution of closed-loop supply chain research. Oper Res 57(1):10–18
7. Li J (2010) The potential for cannibalization of new products sales by remanufactured products. Decis Sci 41(3):547–572
8. Ozkır V, Başlıgıl H (2012) Modelling product-recovery processes in closed-loop supply-chain network design. Int J Prod Res 50(8):2218–2233
9. Subramoniam R, Huisingh D, Chinnam RB (2009) Remanufacturing for the automotive after market-strategic factors: literature review and future research needs. J Cleaner Prod 17(13):1163–1174
10. Yun YS, Chung HS, Moon C (2013) Hybrid genetic algorithm approach for precedence-constrained sequencing problem. Comput Ind Eng 65(1):137–147

Optimization of Supply Chain Based on Internet Plus

Wen Hua(✉)

School of Tourism and Economic Management,
Chengdu University, Chengdu 610000, People's Republic of China
846937934@qq.com

Abstract. "Internet plus supply chain" one of the internet plus strategic action, basing on information technology and platform, promotes the high fusion of elements of supply chain and modern science & technology, and upgrades efficiency of supply chain. The basic factors-information technology guarantee and information sharing, which keep the supply chain operation efficiently, protect the "three flow"-material flow, information flow and fund flow efficient and order movement on the supply chain. This paper analyses the situation and the problem about "three flow", and discusses how to improve the efficiency of supply chain system, from client, to manufacture, and to logistics and financial service with the help of internet technology.

Keywords: Internet plus · Logistics management · Supply chain management

1 Introduction

In 2015, China Premier Li Keqiang puts forwards the "Internet plus" strategic action plan in the government work report to integrate the Internet with traditional industries by means of the Internet platform and information technology, to encourage industrial networks and Internet banking, and to guide Internet-based companies to increase their presence in the international market.

The internet provides information and communication technology and platform for enterprises, promotes the effective communication and coordination between enterprises, business partners, customers and inner-enterprise, and makes it possible to implement supply chain effectively. The enterprises in supply chain can make the synergistic responds quickly, reduce costs of supply chain effectively and improve enterprise efficiency. Even The ratio of logistics cost to GDP keeps declining from 18.4% in 2007 to 16.0% in 2015, and is 0.6% lower than the year earlier [8], it is still higher. The logistics cost in our country is 5% higher than developed countries, which shows the obvious gap. Global Logistics Performance Index (LPI) issued by the World Bank every two years is a international index that demonstrates the domestic logistics abilities of one country or

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_80

one region to participate global supply chain. The 2016 assessment result shows China ranks at 27^{th} [9] and is in the second echelon. As a whole, the status quo of the weak soft power in circulation filed, such as high cost, low efficiency, low intensive level, insufficient industrial support and issues of trust, standard, talents, safe and environmental protection, constraints the efficient operation of supply chain.

The supply chain is a sequential and interconnected organization structure including the design, production and sales and a value chain involving suppliers, manufacturers, transportation and all other components. Meanwhile, the supply chain links business activities to achieve accurate forecast, less inventory and meet real-time needs based on information system. This paper starts with the situation and the problem about "three flow", and discusses how to improve the efficiency of supply chain system, from client, to manufacture, and to logistics and financial service with the help of internet technology.

2 Current Situation of Logistics, Information Flow and Fund Flow in the Supply Chain in China

In the Internet plus era, the supply chain is showing a trend of flat model to ensure that logistics, information flow and fund flow perform effectively and orderly in the system. As globalization lengthens and broadens the supply chain, the competition is more and more intense and the competition between enterprises becomes the competition between supply chains.

2.1 Logistics in Supply Chain

Modern logistics links every supply chain activity from raw material procurement to retails, achieving external integration and reflecting integration advantage. With the speeding up of globalization, raw material procurement and product sales expand in scale in China's manufacturing enterprises. Procurement and sales are attached with the same importance. In terms of transportation, owing to China's road especially expressway construction, consummation and construction of highway networks and transformation of consumption patterns, the road transportation scope was widen and highway logistics accounted for 75.52% of the total freight volume (2015) as shown in Table 1.

Table 1. Annual traffic volume of transportation (10000 ton)

Year	Railway	Highway	Water transportation	Civil aviation	Total
1990	150681	724040	80094	37.0	970602
2000	178581	1038813	122391	196.7	1358682
2015	336000	3150000	614000	625.3	4171000

Source: National Bureau of Statistics et al., China Statistics Yearbook

The data has shown [3] that: In China, there exist over 7 million logistics companies with more than 16 million vehicles and an average of 2 cars for each company. The trucks empty driving rate is 50% and the cargo are transferred 5 to 6 times before transport drivers. As the individual drivers owe more than 90% of road transport capacity, the road logistics has low degree in scale and intensification and road transport is in small, scattered, chaotic, disorder state. All in all, long-distance road transportation has no advantage in economy and environmental protection.

Revenue composition of logistics enterprises: transport and warehousing business make up for 60%, freight forwarding business for 13.9%, distribution and distribution processing for 5% [2], indicating that logistics enterprises focus on transport and warehousing that are the base business with low profit margins. And for profit margin: integrated logistics enterprises accounts for 8.9%, warehousing enterprises for 6.9% and transportation enterprises for 3.3%.

In general, for the majority of small and medium-sized logistics enterprises, they are lack of fund, labor force, capacity and application of information technology; the labor-intensive operation relies on a large number of labors; they have insufficient capacity to provide value-added logistics service since their business focus are basic logistics services such as transportation, warehousing. What's more, the general low profitability can be attributed to two reasons. One is that demand of logistics services reduces as China's economy slows down. At the same time, logistics enterprises has weak ability to provide quality service and low efficiency, which cannot meet the current market demand or meet in the way the market needs.

2.2 The Information Flow in Supply Chain

The value of information technology application and information sharing among upstream and downstream enterprises in supply chain is widely concerned and recognized by scholars and enterprises. From 2013, the National Bureau of Statistics released information technology application and e-commerce situation according to enterprise types, indicating that the application of information technology has increasing impact on China's economy. Table 2 suggests the enterprise information and e-commerce situation in 2014. The proportion of enterprises with websites in manufacturing and transportation industry as well as warehousing industry were separately 70% and 45%, but enterprises dealing with e-commerce accounted for only 8.6% and 4%. The website is mainly used for non-transactional corporate propaganda and product display.

Logistics information technology is divided into four categories: data acquisition technology (such as bar code, RFID technology), information technology, warehousing management technology and transportation management technology, and respectively the proportion of technology used by Chinese logistics enterprises is 56%, 94.9%, 73.8%, 88.5% [7]. According to the market research report on China's logistics industry in 2015, the utilization rate of logistics applications are as follows: storage management system accounts for 17.21%, transportation management system for 17.21%, and vehicle tracking system for

Table 2. Information technology application and E-commerce of enterprises in 2014

Project	Number of enterprises	Number of computers used per 100 workers	Number of enterprises with websites	Number of websites owned by per 100 enterprises	Dealing with E-commerce	
					Number of enterprises	Percentage (%)
Manufacturing	347376	18	244822	70	29938	8.6
Transportation, warehousing, postal service	35187	25	15908	45	1402	4

Source: National Bureau of Statistics et al., China Statistics Yearbook (http://www.stats.gov.cn/tjsj/ndsj/2015/indexch.htm)

19.35% [3]. Application of logistics information technology and the software does not mean that the supply chain has achieved information sharing, collaborative response and quick response. Only by breaking the blocked business model and interest chain established by the system and institution can we cross the boundaries of organizations, regions and industries.

2.3 Fund Flow in Supply Chain

On the whole, the financing difficulties are the bottleneck of enterprise development for many SMEs. With the concept of supply chain management is strengthen, the overall foundation of the supply chain is increasingly apparent. This development has gone through three stages. The first stage, logistics' factors combine with financial activities. The second stage, the development of supply chain finance moves from factors to integrating management of logistics operation, business operation and financial management. The third stage, internet + supply chain finance relieve information asymmetry and reduce the costs of information access and information processing, and the industry supply chains are optimized through financial resources [10]. The Supply chain finance takes the core enterprise as the starting point, and the systematically arranges financing for all member enterprises in the supply chain. Supply chain finance inputs the funds into the small and medium-sized upstream and downstream enterprises in the supply chain, which provides the financial support for the supply chain. The application of Internet achieves financial support for supply and demand, that is, the third-party payment platform reduces financing costs for enterprises in supply chain, and consequently the bottleneck of funding is alleviated.

3 The Optimization of Supply Chain

The essence of "Internet plus" is to online and digitalize the traditional industry. When the commodity display and transaction behaviors shift to Internet and online, information and data can flow, exchange, mine, transfer and share. Besides, supply and demand can be connected at the lowest cost. However, the enormous information and data are featured with scattered property, low value intense, low connection. In this case, big data, cloud storage, data mining and other technologies are adopted to provide valuable information for decision-making by data integration, processing and analysis.

The integration of Internet and industry moves forwards generally from the downstream to the upstream in the supply chain and changes consumers in the first place. According to the Internet Statistical Report of 2015, the popularity rate of Chinese netizens is 50.3%, among which there are 413 million, that 60.03% of the netizens [4] making up for 30.20% of the total population conduct online shopping and transactions (26.68% in 2014). Reports [1] have shown that China's online retail market has 3.8 trillion yuan volume, which accounts for 12.7% of total retail sales(10.6% in 2014), represent the proportion of retails that applying Internet. The application rate of E-commerce, from the consumer → retail → production → transport services along the supply chain, are respectively 26.68%, 10.6%, 8.6%, 4.0% in 2014 as shown in Fig. 1.

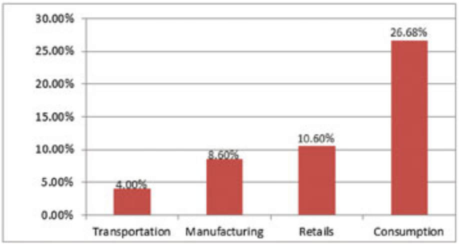


Fig. 1. Application rate of E-commerce in different links of supply chain (Year 2014)

3.1 Demand Side Management-Define and Expand Effective Demands

In the open network environment, the demand of the increasing amount of end users participating in innovation and revolution must be given prominent attention by the supply side and placed in the logic starting point and attention center of enterprise decision making. Otherwise, once the consumers' upgrading needs are not responded, then the low-cost, convenient external resources are their alternatives, which results in decrease of internal demand.

(1) Demand side management-Define and expand effective demands
Diversification and uncertainty of demand reduce the versatility of the product and increase the inventory of supply side and make production and management more difficult. Thanks to the application of Internet information technology, it reduces the cost of collecting, integrating and utilizing consumer information, thus the enterprise is sensitive to marketplace and explores and meets the end-user demand in more convenient, fast, efficient and low cost way. What's more,

business to customer model achieves disintermediation effect, which greatly weakened the profits divided by the distributors, retailers and other traditional business roles in the supply chain and is conducive to get access to first hand market information because supply chain is directed to end users. Data analysis, cloud computing and other tools utilize the multi-dimensional information connect by the internal and external network platform to dock, integrate, analysis trend. Enterprises are enabled to increase supply chain forecasting and decision-making capabilities and conduct scientific regulation of production, inventory, logistics and other key activities in supply chain.

(2) Exploit and foster consumer market by innovative activities
Efficient supply chain must be the interaction, feedback and balance between efficient supply and demand. The supply side should strengthen understanding and awareness of the market demand and positive response, maintaining this market demand is true, reliable and sustainable. Supply chain finance provides credit support for a large number of SMEs in the supply chain, and consequently the consumer finance of supply chain has emerged. The typical representative such as Alibaba, Jingdong Mall and Suning use their own funds to provide online supply chain financial services, among which installment and other shopping financing services increase the consumption by 30%–50%.

By fully utilizing these trading platforms, supply side sells in domestic and overseas markets as well as expands sales channels at a lower cost. Owing to the strategic cooperation with the e-commerce, large-scale production enterprises combine the market development, capital operation and brand upgrading together, improve the ability to meet consumption experience based on Internet and actively respond to market demand for quality products and services. These efforts made by supply side improve brand awareness and acceptance and consequently improve the company's position in the value chain. What's more, the global business opportunities in supply chain are added.

(3) Guide the consumers to deeply participate in supply chain activities
The emergence of E-commerce has changed the objects of orders and logistics delivery from retailers to consumers. Since business opportunities are found from the user's selection and behavior changes, production and service enterprises should attach importance to user experience and customer voice to ensure consumers have more opportunities to participate in the supply chain activities. Consumers become a crucial factor in supply chain management. On one hand, the demand side influences the product and service design of supply chain through order, feedback, reviews and so on. On the other hand, the supply side can effectively analyze preference trends, behavior characteristics and potential demands of consumers through large data analysis, and accordingly make customized designs based on personalized needs. It means that by application of Internet technology, the ideas of individual end consumer become reality products and services. The customized design innovation featuring "User-Generated Content" is becoming the big trends [6].

The visual distribution of E-commerce Suning, a kind of offline services assisted with the "Internet plus", brings great fun for users and enhance the user experience by giving a sense of involvement and control. Now a lot of outstanding enterprises are marching in deepening their involvement in the supply chain management. They not only collect customer information, but also focus on guiding consumers to participate in collaborative intelligent development of product development, customized products, and experience processes. Enterprises are expected to improve the ability to satisfy customer needs based on the Internet and enhance the quality of supply capacity.

3.2 Supply Side Management-Improve Quality Supply

Supply and demand coexist. The supply side reform means that we shouldn't only focus on demand stimulation and it should be emphasized that the economic growth can't simply depend on stimulating the demand. In the long run, the main supporting factor lies in effective supply's response to high-quality demand and guidance. (Jia Kang, 2015) [5] While, the supply side reform should utilize the production factors including the labor, land, resources, capital, technology to achieve the intensive production model characterized by economic innovation and efficiency, and to achieve the upgrading from low-level supply to high-level supply through knowledge, technology, products and institutional innovation.

(1) Data Analysis Transformation-Big Data Analysis
With the speeding up of convenient Internet and wide application of new technologies, the consumption patterns of end-users have changed and demand expectations are becoming diversified and personalized. In this case, manufacturing enterprises should quickly respond to external changes and effectively mine the data potential in order to solve supply chain management issues such as demand forecasting, replenishment forecasting, inventory optimization, foundry collaboration, production arrangements and supplier management. Besides, the manufacturing enterprises should monitor and manage the process and results including process tracking, exception reminder, order tracking, transaction history data and so on. The solution is the big data analysis which is characterized by dynamic real-time, multi-dimension, public sharing and low cost. Based on the connection of data, big data can help all the node enterprises in the supply chain to watch and control every activity including product design, purchasing, manufacturing, order processing, logistics, financial services and consumers. Besides, it can display inventory, production, orders, circulation, distribution and other activities in real time, and assist the enterprises to optimize, forecast and adjust the balance of supply and demand.

The production systems can cope with the uncertainty brought by the rapidly changing environment. By application of information technology and Internet, the production system can do as follows: ① Grasp the true needs of the market. Process the fragmented information through big data, cloud computing and other information technology, thus the demand trends will assist decision-making; ② Improve production flexibility. It means to promote the production management that quickly responds to multiple products, small batch and rapidity. The "rigid" production mode of large-scale production transforms to the adaptable and "flexible" production mode; ③ Integrate external advantage resources timely.

The enterprises in supply chain can share information at a lower cost through the information platform.

(2) Manufacturing and production transformation-Quick response
O2O model, characterized by small batch, multiple species and fast response, has become a popular market demand. The "Internet plus" mode accelerates this consumption and further improve the consumer expectations. The changes are so fast that manufacturing enterprises are confronted with the following situations: ① Due to poor versatility of personalized products, inventory risk is greatly increased. ② As the technology and product updates speed up, the marketability value of the inventory and performance are reduced, bringing huge economic losses. Therefore, it is necessary to adjust the established production management, system arrangement, equipment selection, logistics and warehousing business pattern that targeted for large-scale production, instead, to transform to flexible manufacturing by full application of Internet technology.

From perspective of production management, it is necessary to adjust the production arrangement timely, quickly and frequently and adopt a flexible and diversified combination of production factors. Besides make full use of internal and external resources advantages to prevent excessive production and meet the market demand at the same time. From the perspective of production equipment, various artificial intelligence tools, manufacturing equipment and computing methods are widely used in the manufacturing process. Manufacturing intelligence is fully utilized in the scheduling, design, processing, control, planning and other aspects of the production. Manufacturing enterprises must comprehensively utilize manufacturing technology, management science, computer science, information technology, and conduct comprehensive processing of internal and external information and reasonably schedule the procurement, logistics and production. The flexible production system based on information technology can respond to the changing production and sales, and reduce response time to changes, besides, production costs can fluctuate with the amount fluctuation.

(3) Structural transformation of supply side-Provide quality supply
On one hand, the overcapacity is prominent when shortage economy is shifting towards surplus economy in China. On the other hand, as Chinese per capital income continues to grow, the basic needs upgrade to development demands and the demands are diversified and personalized. In the Internet era, consumers have more choices in products, service, consumption channels and manners. The popularity of overseas purchase heat demonstrates that people welcome continuous protected products of high quality and fair, just and honest service. When consumers are dissatisfied with the products and services, they turn to lower cost channels as alternatives. Therefore, the effective way to resolve overcapacity and overstocked products is to change the export-oriented supply system to improve the effective domestic demand, provide favorable environment for domestic high-quality supply enterprises to ensure supply keep ups with the pace of demand.

Expanding domestic demand has always been an important means of China's economic development. In practice, domestic demand growth comes mainly from investment as the contribution of terminal consumers is limited. Supply side structural reform, in essence, is the effective supply and utilization of production factors. In particular, it gives full play to the role of capital and technology factors. It means we should introduce advanced management ideas, improve management capacity, reduce costs, eliminate backward production capacity, get rid of low-cost competition, improve supply quality, and create space for the upgrading of industry and supply chain.

To start with the supply side, it is necessary to resolve the contradiction between consumer demands for quality products and services and the existing low-end products provided by production and service enterprises. "Internet plus" information technology reallocates the resources and provide a strong platform to eliminate the information asymmetry caused by mismatched contradictions.

3.3 Enhance Logistics and Financial Services Capabilities

The modern logistics industry provides material flow service for collaborative activities of enterprise in supply chain node, and improves the interconnection and interdependence among the supply chain nodes. The financial innovation in supply chain on this ground provides the financial support for supply chain development. However, the high financing and logistics costs have restricted the competitiveness of China's industries, affecting the efficiency of the entire industrial structure. Therefore, China's supply side reform not only involves the production field, but also includes the reform on the transaction behaviors, systems and circulation and service areas among the manufacturers, agents, dealers, retailers in the supply chain.

(1) Integrate the fragmented logistics services and improve logistics supply capacity and level
The development direction for the logistics industry to enter into network and organized service is to integrate resources, expand logistics enterprises scale, improve the introduction and application of information technology to provide one-stop service for consumers, expand the logistics services scope and provide value-added logistics services. The logistics platform gathering large sums of fund and technology will become the breakthrough of logistics industry integration.

Various forms of logistics network platform coordinate the goods and capacity, integrate scattered transport resources and sources of goods and improve the efficiency of material circulation. When the supply chain efficiency is low, the supply of raw materials cannot make timely production. Then most of the manufacturers will prepare front-end raw material storage and back-end finished products storage, it might cause high inventory costs and overstock inventory. With the development of the logistics platform and popularization of smart phones and networks, the highway logistics APPs have emerged. These APPs help obtain real-time logistics information conveniently, breaks information asymmetry, matches freight resources effectively and reduce resources wasting. Many small micro-logistics enterprises link the supply side and demand side through the supply chain. With the supply chain information platform, theses companies can rapidly accumulate massive data, accurately research into demand and participate in supply chain services in real time.

(2) Overall penetration of financial innovation on "Internet plus"
To provide financial guarantee for the supply side reform, it is necessary to stimulate innovation through policies, improve the total factor productivity, rationalize the imbalance of supply and demand structure, make the capital flow to the real economy and optimize the allocation of financial resources. Financial innovation should have all-round penetration on "Internet+" in the economy and living. Supply chain finance means that regarding with the real business context and the credit level of the core enterprises of the supply chain, the financial institution provides credit support to small and medium-sized enterprises in the upstream supply chain and consumer finance for downstream consumers. Since core enterprises are the backbones of the supply chain system, they must establish an efficient resource sharing system to provide a sharing platform, maintain membership cooperation, find value creation methods and finally improve the competitiveness of the whole supply chain system.

4 Conclusion

The wide coverage and availability of the Internet has transformed the supply chain from a single, linear, discontinuous information flow into a complex information network, which increases the possibility of choosing excellent supply chain partners and expands the access of quality products and service for consumers. The impact of "Internet plus" on the supply chain, on one hand, is reflected in the application of information technology and big data, by which companies achieve real-time information, data collection and analysis, resulting that business decision-making is no longer simply based on structured data; on the other hand, the impact is reflected in the changes of production mode and the circulation ways. The Internet plus promotes the development of multi-industry convergence and transformation and upgrading of traditional industries, and fundamentally reduces the contradictions between high inventory and low consumption satisfaction as well as between overcapacity and sluggish manufacturing.

The transformation process from the traditional supply chain management to the "Internet plus supply chain" is not the simple application of information technology to supply chain management, but the conversion and innovation of information technology from subordinate roles to the dominant rules, which promote customers' true demands capturing, supply chain visualization process, agile manufacturing, real-time data acquisition and sharing, and partner collaboration, in the environment of continuous technology development of internet, big data and cloud computing. A reviewed way of thinking, about market, custom, production and enterprise value even the business ecosystem, is revealed.

Acknowledgements. Sichuan Province Key Research Base of Social Science—Sichuan Province Cycle Economy Research Center, "Ability Evaluation Research of the County Economy Development from the Perspective of Cycle Economy" (XHJJ-1610); Sichuan Province Key Research Base of Humanities and Social Sciences–Sichuan Province Education Development Research Center of China West Normal University, "Adaptability Research between Higher Education Subject Distribution and Industry Cluster in Sichuan Province" (CJF16007).

References

1. CEBRC (2016) Online retail market data monitoring report in China in 2014 [DB/OL]. China E-Business Research Centre. http://www.100ec.cn
2. CFLP (2016) statistical investigation report on logistics of key enterprises in China in 2015 [DB/OL]. China federation of logistics & purchasing, China logistics information center. http://www.clic.org.cn/wlxxwlyx/262043.jhtml
3. CFLP (2015) China logistics development report (2014–2015). China Federation of Logistics & Purchasing. China Logistics Association. China Fortune Press, Beijing
4. CINIC (2016) Statistics on the development of the 37th China internet network [DB/OL]. China Internet Network Information Centre. http://www.clic.org.cn/wlxxwlyx/262043.jhtml
5. Jia K, Shu J (2015) Economic 'new framwork' and 'new supply': the pursuit of important link and epitomize in innovation. Sci Technol Prog Policy 11:8–14
6. Li J, Liu C, Zhang N (2016) Anlyzing of design process and operational mode of different crowd-sourcing supply chains with internet plus. Sci Technol Prog Policy 11:24–31
7. Lin CY (2008) Determinants of the adoption of technological innovations by logistics service providers in China. Int J Technol Manage Sustain Dev 7(1):19–38
8. NDRC (2016) The national logistics operation status briefing in 2015 [DB/OL]. National development and reform commission. http://www.sdpc.gov.cn/jjxsfx/201605/t20160531_806056.html
9. NDRC (2016) The world bank logistics performance index in 2016 [DB/OL]. China Logistics Information Center. http://www.clic.org.cn/kjwl/270223.jhtml
10. Song H, Chen SJ (2016) Development of supply chain finance and internet supply chain finance: a theoretical framework. J Renmin Univ China 35:10–16

Assessing the Recycling Efficiency of Resource in E-Waste Based on MFA of Reverse Logistics System

Minxi Wang, Xiaoling You, and Xin Li(✉)

College of Management Science, Chengdu University of Technology, Chengdu, The People's Republic of China
Lixin2012@mail.cdut.edu.cn

Abstract. The waste electrical and electronic equipment (WEEE, e-waste) has caused many issues, such as waste of resource and environmental pollution. This paper applied material flow analysis (MFA) method to analyze the flows of recycled resource from WEEE and process of reverse logistics from recycling to reuse. And analyzed the discarded TV data in K company (a leading television producer in China), and calculated the recycling efficiency of metals. The research showed that recycling efficiency of the metals from discarded TV in K company is 12.3%, which is far below the criterion (In Japan, the "Household appliance recycling and reuse" recycling efficiency of useful resource from discarded television by the television manufacturers must be above 55%). At last, we gave some suggestions for developing recycling efficiency of manufacturing enterprises and recovery enterprises in China: (1) Improving the recycling efficiency of manufacturer, retailer and consumer can contribute to improve the recycling efficiency in the whole reverse logistics system; (2) Improving the technological level of decomposition and dismantling of e-waste in the recycling process can increase the recycled material; (3) Manufacturing enterprises should have the green production concept and the participants in the supply chain should follow the concept of EPR (Extended Product Responsibility).

Keywords: Reverse logistics · Recycling efficiency · MFA · WEEE

1 Introduction

WEEE is the waste of electrical or electronic products, it covers a wide spectrum ranging from consumer goods such as discarded refrigerators, air conditioners, washing machines, televisions and computers, mobile phones to capital goods such as some unqualified products, parts and scraps in the production and repair processes [6].

Because of containing precious metals (such as gold, copper, aluminum and silver) and plastic, the reutilization of WEEE has magnificent recovery value and

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_81

environmental protection advantages. Recycling resource from the WEEE has long-term strategic significance, not only can reduce the waste of resource, promote social development and regional economic growth, but also can reduce the serious threat from toxic and harmful substances in e-waste to the environment and human health.

Many researchers have made a large amount of researches on the recycling in reverse logistics system and e-waste recycling. In the aspect of reverse logistics system mainly focused on architecture design for the network, model establishing and waste management in reverse logistics system. Reverse logistics network can be divided into different branches according to the type of goods in reverse logistics system, applying the mixed-integer linear program (MILP) or integrating the analytical hierarchical process (AHP) can address the complex network configuration of a reverse logistics system [2,9,11]. In order to design the network of reverse logistics system, the capacities of each nodes and the number of recycling facilities should be considered. By applying some models can minimize the costs, and simultaneously can consider green and social issues to design a reverse logistics network [10]. In addition, establishing the model of reverse logistics also can solve some practical problems, for instance Ehab Bazan established the model based on the economic order/production quantity (EOQ/EPQ) and the joint economic lot size (JELS) settings for reverse logistics inventory systems [5]. Waste management is an important part in reverse logistics system, and waste recycling is a multi-disciplinary problem, a holistic view and viewpoints from different decision levels should be considered when modelling a reverse logistics system for waste management [7]. Setting up the waste management strategies can help firms to reduce waste and have a healthy reverse logistics system [15]. In the aspect of e-waste recycling, developing model and applying method can analyze the WEEE based on a real world case. For instance, Ayvaz, B created the two stage stochastic programming model provides acceptable solutions to make efficient decisions for reverse logistics network design of e-waste [4]. By applying the material flow analysis method can make a detailed analysis for recycling and stock of e-waste, as well as the recycling efficiency assessment and optimization in the closed-loop supply chain [1,14].

This paper combine recovery system with material flow analysis method in reverse logistics system and analyze the process of resource recycling of e-waste. Take recovery of discarded televisions of K company (a leading television producer in China) as an example in this thesis, its recycling efficiency of metal in the discarded television is calculated and assessed, based on that we gave some suggestions for the manufacturing enterprises and recovery enterprises in China.

2 Theory and Methodology

2.1 Reverse Logistics and Reverse Logistics Network

A complete definition of reverse logistics is put forward by the European Working Group on Reverse Logistics, and is formulated as follows: "The process of planning, implementing and controlling backward flows of raw materials, in-process inventory, packaging and finished goods, from a manufacturing, distribution or use node, to a node of recovery or node of proper disposal" [3].

Reverse logistics network is the nodes distribution and the arrangement of transportation routes between each node in the reverse logistics system. Typical reverse logistics network is shown in Fig. 1.

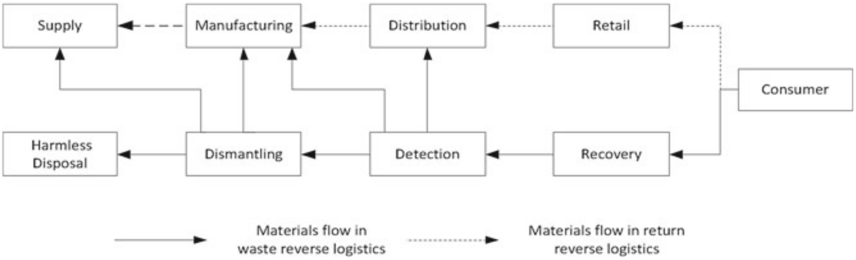


Fig. 1. Typical reverse logistics network

2.2 Material Flow Analysis

Material flow analysis (MFA) is "a systematic assessment of the flows and stocks of materials within a system predefined in space and time". The method mainly

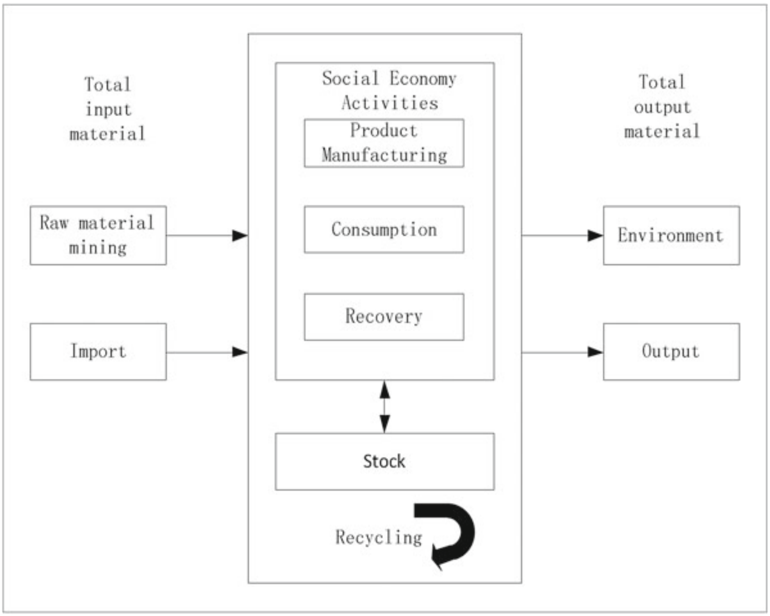


Fig. 2. The basic framework of material flow analysis method

involved the material inputs, outputs, and inventories [8]. According to the law of conservation of mass, MFA could get an ultimate conclusion that indicates the status of material input, stock and output in each stage of life cycle. It is based on the material balance theory (Total material input = Total material output + Stock), by means of material flow analysis, we can estimate and evaluate the condition of development, utilization and waste management of natural resources and various substances. The basic framework of material flow analysis method is shown in Fig. 2.

2.3 Material Flow Analysis of Reverse Logistics System

MFA of reverse logistics system is the main methodology in this paper which is a way to analyze the material flow of e-waste in reverse logistics system.

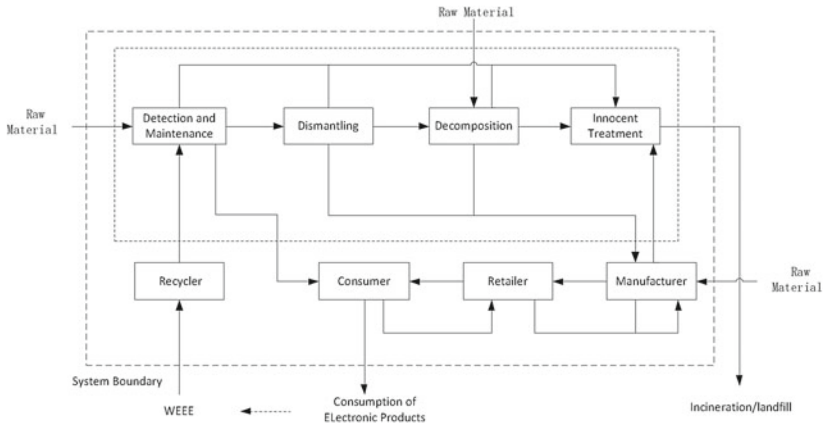


Fig. 3. The material flow analysis frame of the reverse logistics system

As shown in Fig. 3, participants in the frame are recycler (which mainly responsible for four recycling process: detection and maintenance, dismantling, decomposition and innocent treatment), manufacturer, retailer and consumer.

3 Analysis and Discussion

3.1 Reverse Logistics Process for the WEEE

As a special kind of reverse logistics, the recovery process for e-waste includes five aspects. Figure 4 shows the principal stages in the recovery process.

(1) Recovery

Recovery phase is the first and most important stage in the whole recovery process. At present, the main sources and recovery channels of e-waste in China are shown in Table 1.

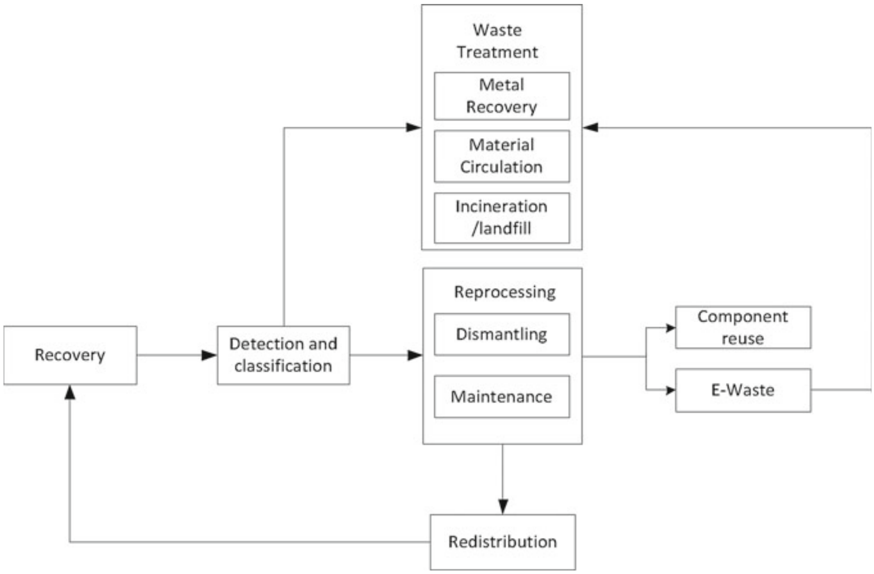


Fig. 4. Recovery process for the WEEE

Table 1. The main sources and recovery channels for e-waste

Category	Comment
Main source	The daily life of the residents; Electronic products and household appliances manufacturers and sellers; Illegal imports of e-waste from abroad.
Recovery channel	Door-to-door recovery; For trading-in old one for new one in the Shop; Flea Market of electronic products and household appliances; Garbage recycling by government/The glean and collect scrap man; After-service station

(2) Detection and Classification

E-waste has a wide variety of categories, and each category can be divided into small classes and the corresponding value of it is different. Therefore, the work that must be done is rigorous detection and classification for the recycling of e-waste according to their characteristics such as structure or components function to determine the subsequent process, such as after maintenance to resale, reuse the parts after dismantling, material recycling after decomposition or waste treatment.

(3) Reprocessing

In the reprocessing stage, e-waste can be reacquired the value that reuse, remanufacture or recycle through reasonable reprocessing for recycled products or parts. Reuse only aims at the recycled products can be used directly through cleaning or simple maintenance, such as the toner cartridge which can be used again through the simple work. Remanufacturing aims at the e-waste can enter the manufacturing stage again after dismantling, replace or repair. And the recycle aims at the parts, glass and plastic in the recycled e-waste.

(4) Redistribution

The recycled e-waste can be as commodity to reuse and enter the consumer market or donate to the consumers in poor areas directly after the inspection and reprocessing stage. In this stage, distribution is the most important work, which can make the entire recovery process for the WEEE operate efficiently.

(5) Waste Treatment

For the e-waste that has no circular economic value or can bring large harm to the environment, the available material in them can as the recycling resources to be reused through the reasonable treatment, such as dismantling, melting, refining and electrolysis. While for the e-waste cannot be reused, the disposing is partitioned in two ways, one which is stored them in landfills; the other approach is sent them to an incinerator.

3.2 MFA in Reverse Logistics System for E-Waste

(1) Date Acquisition

In this paper, according to the MFA of reverse logistics system, we built a reverse logistics system for e-waste based on the self-operated recovery system of the K television manufacturing company in China. Take the recycling efficiency of the metal resources (copper, aluminum, iron) of the company as the concrete calculation example. Through collecting television sales data of the company over the years and applying the market supply method A [12] to forecast the amount of discarded TV of the K television manufacturing company in 2016.

Through the existing research [13] can get the duration of use and the waste rate of the television is shown in Table 2.

Table 2. The duration of use and the waste rate of the television

Duration of use(year)	< 8	8–9	9–10	10–11	11–12	12–13	13–14	14–15	15–16	> 16
Waste rate	6%	6%	10%	13%	15%	15%	13%	10%	6%	6%

The sales date of K television manufacturing company over the years is shown in Table 3.

The market supply method A formula as follows:

$$Q = \sum S_i \times P_i, \qquad (1)$$

Table 3. Sales of television over the years

Year	Sales (Million)
2000	1.06
2001	1.09
2002	1.63
2003	2.41
2004	2.43
2005	3.99
2006	3.99
2007	3.54
2008	4.01
2009	5.8
2010	6.02
2011	6.41
2012	7.56
2013	8.05
2014	9.67
2015	9.92

where, Q represents the amount of discarded televisions a year, S_i represents the sales of televisions i years ago (assuming all the sales of televisions are used), P_i is the waste rate represents the proportion of the discarded television in sales i years ago.

By using this model, it can be concluded that the amount of waste television in 2016 is 2.12million. Combining with the relevant information from the related industry, 50% of discard television can be recovered, and the dismantling rate (after dismantling can be recycled parts ratio) is set to 35%, and the decomposition and utilization rate is set to 56%. The average weight of a television is 20 kg, the content of copper in a television is 5.4%, the content of aluminum in a television is 5.4%, and the content of iron in a television is 5.3% [18]. Because the resource loss of recycled metal to the environment in each stages, and in order to facilitate the research, so the following assumptions are given:

(1) The structure of resource recycling in reverse logistics system is a closed loop network, all the discarded television after disposing would return to manufacturer;
(2) The return production from consumer must be sent to the manufacturer for processing;
(3) The product could be reused after detection and maintenance would be resold and enter the consumer market;
(4) All the discarded television (which could be recycled) must be recycled by the recycler.

Therefore, the amount of recycled metal from discarded television is 3413.2 tons. The specific data are shown in Fig. 5.

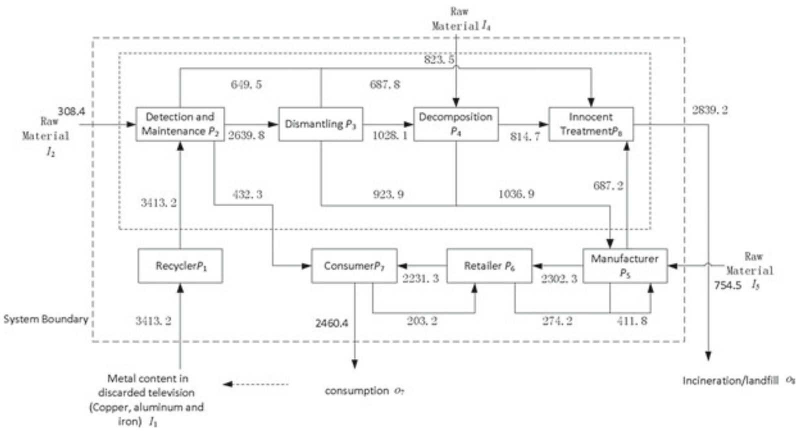


Fig. 5. The specific data in reverse logistics system of discarded television. All units measures in tons

(2) Empirical Analysis

According to the data in Fig. 5, the input-output table of reverse logistics system of discarded television is shown in Table 4.

Table 4. The input-output table. All units measures in tons.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	I_0
P_1	0	0	0	0	0	0	0	0	3413.2
P_2	3413.2	0	0	0	0	0	0	0	308.4
P_3	0	2639.8	0	0	0	0	0	0	0
P_4	0	0	1028.1	0	0	0	0	0	823.5
P_5	0	0	923.9	1036.9	411.8	274.2	0	0	754.5
P_6	0	0	0	0	2302.3	0	203.2	0	0
P_7	0	432.3	0	0	0	2231.3	0	0	0
P_8	0	649.5	687.8	814.7	687.2	0	0	0	0
O_0	0	0	0	0	0	0	2460.4	2839.2	0

In order to quantify the relationship of various stages of metal cycle in reverse logistics system, we give the mathematical expression of input-output relationship based on material balance theory as Eqs. (1) and (2), x_k is the total flow in P_k, f_{kj} is the flow of P_j to P_i.

$$x_k = \sum_{i=1}^{n} f_{kj} + I_k \qquad k = 1, 2, \cdots n, \tag{2}$$

$$x_k = \sum_{i=1}^{n} f_{ik} + O_k \qquad k = 1, 2, \cdots n. \tag{3}$$

According to the Eqs. (2) and (3), the total flow of each node is shown in Table 5. In order to assess the overall recycling efficiency of the system, we give the mathematical expression Eq. (4)–(6).

Table 5. The total flow of each node

Node	Amount (tons)
P1	3413.2
P2	3721.6
P3	2639.8
P4	1851.6
P5	2989.5
P6	2505.5
P7	2663.6
P8	2839.2

$$n_i = \frac{I_i + \sum_{j=1}^{n} f_{ij}}{o_i + \sum_{j=1}^{n} f_{ji}}, \tag{4}$$

$$R_i = \frac{n_i - 1}{n_i}, \tag{5}$$

where, n_i is the relation of the input and the final output of a single node. When $n_i = 1$ indicates that the recycled metal resources through the node i, but not returned to the node i for recycling. When $n_i > 1$ indicates that the recycled metal resources through the node i, and returned to the node for recycling with a direct or indirect way.

n_i is recycling efficiency of a single node. When $R_i = 0$ indicates the flow of the recycled metal resources through the node i is unidirectional. When $R_i > 0$ indicates the recycled metal resources through the node i can return the node for circulating and using.

Table 6. The calculation results of n_i and R_i

	1	2	3	4	5	6	7	8
n_i	1	1	1	1	1.14	1.09	1.19	1
R_i	0	0	0	0	12.30 %	8.30%	16%	0

The calculation results of n_i and R_i are shown in the following Table 6.

According to weighted method, we use the recycling efficiency of each node can get the overall recycling efficiency C in the system.

$$C = \frac{s_c}{s} = \frac{\sum_{i=1}^{n} R_i \times x_i}{\sum_{i=1}^{n} x_i} = \frac{1001.841}{8158.6} \approx 12.3\%, \tag{6}$$

where, S_c is the total recycling flow, S is the overall flow of nodes which the recycling efficiency of a single node is not equal to zero [16].

Therefore, the overall recycling efficiency of the system is 12.3 %. That is the recycled metal resources after the whole recovery process, 12.3% metal resources in discarded television can be reused.

4 Conclusions

According to the analysis and discussion above, we concluded that the recycling efficiency of metal resources in waste television of K television manufacturing company which is representative in the television industry, so the recycling efficiency has a certain reference value for home appliance manufacturing industry in China.

There is not a sound assessment standard to evaluate the recycling efficiency in China, so this paper take Japanese experience as reference. In Japan, the "Household appliance recycling and reuse" started from 2001 specifies that the recycling efficiency of useful resource (metal in the majority) from discarded television by the television manufacturers must be above 55% [17]. Obviously, the recycling efficiency of the television company was hard to reach the criterion with the present resource recycle level, so it is feasible to make optimization for reverse logistics recycling system of e-waste in China, and improve circulation efficiency of the recycling system, so we give the following suggestions:

(1) From the above calculation process, the overall recycling efficiency lever mainly depends on P_5, P_6 and P_7 (Fig. 5). Thus, we can increase the numerical value of P_5, P_6 and P_7 (Table 6) to increase the overall recycling efficiency, that means improving the flow of manufacturers, retailers and consumers is the key problem to improve the recycling efficiency of the system.

(2) Improving the secondary utilization ratio of the material in the recycling system, such as improving the technological level of decomposition and dismantling of e-waste in the recycling process, the recycled material in e-waste can be increased, but the input of raw material can be reduced, and then minimize the wastage of resource from the external environment.

(3) Chinese electronic products and household appliances manufacturing enterprises should have the green production concept to decrease the amount of waste that can't be reused. In addition, the participants in the entire supply chain should follow the concept of EPR (Extended Product Responsibility). That means not just the manufacturer, the distributor, retailer and consumer should also strengthen mutual cooperation and bear their respective responsibilities, for the purpose of promoting the level of reuse and recycling of resources.

Acknowledgements. We would like to thank all anonymous reviewers for their constructive comments and suggestions. This research was supported by the fund of the Excellent Innovation Research Team at Chengdu University of Technology (KYTD201406).

References

1. Agamuthu P, Kasapo P, Nordin NAM (2015) E-waste flow among selected institutions of higher learning using material flow analysis model. Resour Conserv Recycl 105:177–185
2. Alshamsi A, Diabat A (2015) A reverse logistics network design. J Manufact Syst 37:589–598
3. Alumur SA, Nickel S et al (2012) Multi-period reverse logistics network design. Eur J Oper Res 220(1):67–78
4. Ayvaz B, Bolat B, Aydin N (2015) Stochastic reverse logistics network design for waste of electrical and electronic equipment. Resour Conserv Recycl 104:391–404
5. Bazan E, Jaber MY, Zanoni S (2015) A review of mathematical inventory models for reverse logistics and the future of its modeling: an environmental perspective. Appl Math Model 40(5–6):4151–4178
6. Bertram M, Graedel TE et al (2002) The contemporary european copper cycle: waste management subsystem. Ecol Econ 42(1–2):43–57
7. Bing X, Bloemhof JM et al (2015) Research challenges in municipal solid waste logistics management. Waste Manag 48:584–592
8. Brunner PH, Rechberger H (2006) Practical handbook of material flow analysis. Int J Life Cycle Assess 10(9):293–294
9. Galvez D, Rakotondranaivo A et al (2015) Reverse logistics network design for a biogas plant: an approach based on MILP optimization and analytical hierarchical process (AHP). J Manufact Syst 37:616–623
10. Govindan K, Paam P, Abtahi AR (2016) A fuzzy multi-objective optimization model for sustainable reverse logistics network design. Ecol Ind 67:753–768
11. Mangla SK, Govindan K, Luthra S (2016) Critical success factors for reverse logistics in indian industries: a structural model. J Clean Prod 129:608–621
12. Simon W, Noel D, Matt C (2001) Waste from electrical and electronic equipment. US Environmental Protection Agency
13. Song X (2008) Establishment and economic analysis of the standardized recycling system for waste and used home applications. Doctoral dissertation
14. Su LX, Wang ZH et al. (2012) MFA-based resource cycling efficiency analysis for a closed-loop supply chain system. Ind Eng J

15. Van Der Wiel A, Bossink B, Masurel E (2012) Reverse logistics for waste reduction in cradle-to-cradle-oriented firms: waste management strategies in the Dutch metal industry. Int J Technol Manage 60(1/2):96–113
16. Wu Z (2013) Reverse logistics management ł the WEEE as the research objects. Science Press in Beijing
17. Yu P (2012) An analysis of Japan's home appliance recycling. Contemporary Economy of Japan
18. Zeng X (2008) Mechanism and technology of nonferrous metals recycling from typical e-waste parts. Doctoral dissertation

Analysis of Supply Chain Collaboration with Big Data Suppliers Participating in Competition

Shuai Liu and Hongchun Wang(✉)

School of Economics and Management Engineering,
Beijing University of Civil Engineering and Architecture,
Beijing 100044, People's Republic of China
wanghongchun@bucea.edu.cn

Abstract. The combination of big data theory and supply chain management theory has brought the subversive change for the enterprise cooperation mode, the paper considered that when the big data suppliers to participate in the competition of the supply chain, the condition of overall profit of the supply chain in the two cases of cooperation and non-cooperation decision making. Firstly, the profit model of each member enterprise is established by the demand function theory, and the whole profit model of the supply chain is obtained in the case of decentralized decision-making, then, the profit model for common decision making of supply chain based on the maximization of supply chain is analyzed. After compared the two profit models, the conclusion was obtained that when big data suppliers join supply chain competition and one party gains the dominant position in the supply chain, collaborative decision-making is the key to enhancing the overall profitability of the supply chain.

Keywords: Big data · Supply chain collaboration · Competition · Cooperation

1 Introduction

The method of supply chain management based on mutual cooperation between enterprises is one of the most effective business strategies adopted by enterprises in the fierce market competition, and the rise of big data industry is injected new blood on Supply Chain Management. Data analysis methods are applied to all aspects of supply chain management to effectively curb the procurement shortage, lag of logistics, production bottlenecks. There are some significant achievements in the area of big data supply chain management. Firstly, supply chain management has been intelligent, and a new generation of radio frequency identification technology can provide more reliable data for supply chains, these data flow in the supply chain nodes, so that enterprises can exchange information timely, the operating mechanism of the system can be optimized and the decision making become automated. In addition, big data has also brought disruptive changes for the supply chain logistics. At present, enterprises can already

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_82

achieve real-time monitoring of logistics and route optimization in the use of big data technologies. Transport capacity and efficiency have been greatly improved than before. The focus of future research is to predict supply chain behavior rely on large data. Accurate forecasting is one of the effective measures to help enterprises save costs, and the accuracy of large data projections is not reached by other methods.

Current research on competition and cooperation model of the supply chain is very mature, Yang and Bai, who build up on the Hotelling model to analyze profit-sharing contracts between the supplier and the core firm, and study the impact of upward decentralization on the core firm and the whole supply chain under the assumption that the supplier does not have any advantage on production costs [12]. Wu and Han, who researched the optimal production decisions in closed-loop supply chains consisted of two competing manufactures and one retailer with remanufacturing costs disruptions [9]. Xu et al. [11] gave a literature review of supply chain contracts selection and coordination based on the supply chain under competing multi-manufactures. Shi, who studied in non-cooperative static game and Stackelberg game optimal advertising strategy in supply chain under master-slave, proves that relevance of the manufacturer and the retailer's own ads level and cost and overall supply chain advertising level and costs [3]. Strategic of partner selection of supply chain model was constructed by Xin and others based on game theory, and get optimal strategies for select partners in the case of incomplete information [10]. Zhang, who thought the whole supply chain benefits and members of associated enterprise level knowledge sharing is very important, In light of this, he builds supply chain profit distribution model based on knowledge-sharing cost [14]. Chen et al. [1] proposed the question of how "big data analysis affects the value creation of the supply chain", and it was answered through the use of dynamic capability theory as a unique information processing capability.

Research on supply chain based on the theory of large scale data, Wu [8] based on dynamic differential game theory constructed the dynamic corporation model of three-level supply chain under retailer paid contract, union pay contracts, with big data service providers to participate in, and discusses contract parameters on the profit impact of supply chain. Wang and Cong, [7] who studied the cloud computing service providers involved in two-stage supply chain coordination problem and introduced in the competition model of supply chain revenue-sharing contract and punished the contract so as to verify that the supply chain coordination and decision. Waller [5] established a matrix model to solve when it should be applied or avoid data analysis, theory of how to use predictive analytics improve issues such as promoting lower total logistics costs. Tan [4] through impedimented to data acquisition analysis, set up a corresponding data infrastructure framework, the framework of internal data, available through collection of deep mining, collection capacity required data to build a data network and helped provide supply chain optimal decision. A systematic framework for the SAM (Roadmap, Strategic Consolidation, and Assessment) roadmap was roadmap to help companies avoid getting caught in a mess of data by Sanders [2]. And the concept of supply chain analysis (SCA) driven by big data was proposed and the function, process-based, collaborative and agile SCA Maturity Mode was established by Wang and Angappa [6].

This paper analyzes the overall profitability of the supply chain under different decision-making situations when the big data suppliers participate in the supply chain competition. The supplier and manufacturer's demand curve is taken as the input to construct the profit model of the supplier and the manufacturer when the manufacturer assumes the big data cost respectively. Through the analysis of the model, the profit model of the supplier and the manufacturer is obtained respectively in the decision - The overall profitability of the supply chain. After comparing the two profit models, we get the corresponding conclusion.

The rest of the paper is organized as follows. In Sect. 2, the problem of is described and some assumptions is proposed. In Sect. 3, the model of decentralized decision-making profit and common decision making are built. In the Sect. 4, two models are compared and analyzed. And there is a numerical analysis in Sect. 5. In Sect. 6, conclusions are given.

2 Problem Description and Model Assumptions

The paper considers the supply chain system as a kind of supply and demand relationship between a supplier and a manufacturer. The manufacturer produces a product with a fixed life cycle and mature market demand. The supplier provides raw materials. The input-output ratio of the raw material to the final product is 1: 1. In order to solve the problem of lagging information transmission among member firms and to improve the accuracy of market demand forecasting, the third-party big data enterprise is used to help the supply chain to build a big data decision platform. The cost is borne entirely by the manufacturer. Because of exists of the supply chain big data platform, the information on the supply chain is completely symmetrical. Assuming that the profit-sharing problem in cooperative decision-making has already reached a corresponding contract between supply and demand, the paper will not consider it.

Suppose that the manufacturer's market demand function is $D(p_A) = a - bP_A + bK(t)$, then we can get the manufacturer's price curve $P_A = \frac{a}{b} - \frac{1}{b}D$, for the convenience of calculation, let $\frac{a}{b} = m$, $\frac{1}{b} = \mathrm{n}$, then there is $P_A = m - nD + K(t)$. It is worth noting that, where m is the limit price that the market can afford [13], that is to say, at this price, the purchase rate of the commodity is 0; Suppose that the market demand function of the suppliers is $D(p_B) = \alpha - \beta P_B$, we can get $P_B = \frac{\alpha}{\beta} - \frac{1}{\beta}D$, let $\frac{\alpha}{\beta} = s$, $\frac{1}{\beta} = t$, then $P_B = s - tD$, empathy, s is the limit price that the market can afford.

The symbols involved in the article are described as follows:

D : Market demand;
P_A : The price of the product produced by the manufacturer;
P_B : The price of the raw materials supplied by the supplier;
G_B : The prices of the raw materials supplied by the sub-suppliers;
$K(t)$: Big data horizontal curve;
r : Big data cost coefficient;
Φ_A : The profits of the manufacturers;
Φ_B : The profits of the suppliers;
Φ_{AB} : The total profit of the supply chain when making decisions together

Among them, the big data cost is $F = \frac{rt}{2}D$, that is, the manufacturer's output is positively correlated with the big data costs, the higher the output, the greater the cost of the big data.

3 Building of the Model

3.1 Decentralized Decision-Making Profit Model

In the case of decentralized decision making, the manufacturer's profit depends on the market price of its own products and the price of raw materials from the suppliers, In addition, because of the cost of big data suppliers to join the supply chain is entirely borne by the manufacturer, so the profits are:

$$\Phi_A = (P_A - P_B) \times D - \frac{rt}{2}D. \tag{1}$$

The profit of the suppliers depends on the price of the product to the manufacturer and the purchase price of the raw material. Therefore, the profit of the suppliers is:

$$\Phi_B = (P_B - G_B) \times D. \tag{2}$$

It can be known from the supplier's demand function that $D = \frac{s - P_B}{t}$. Therefore, its profit is:

$$\Phi_B = \frac{1}{t}\left[s\left(P_B - G_B\right) - {P_B}^2 + P_B G_B\right].$$

Find the partial derivative of P_B in the supplier profit model and let it be equal to 0, and then the function of raw material price provided by the suppliers and cost is obtained:

$$\frac{\partial \Phi_B}{\partial P_B} = \frac{1}{t}\left[s - 2P_B + G_B\right] = 0,$$
$$P_B = \frac{s + G_B}{2}. \tag{3}$$

The manufacturer's profit formula can be known from formulas (3) and (1):

$$\Phi_A = [(m - nD + k(t) - \frac{s + G_B}{2})] \times D - \frac{rt}{2}D.$$

Find the partial derivative of D and make it equal to 0, the manufacturer's optimal production decision can be obtained:

$$\frac{\partial \Phi_A}{\partial D} = (m - 2n + k(t) - \frac{s + G_B}{2}) - \frac{rt}{2} = 0,$$
$$D = \frac{m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2}}{2n}. \tag{4}$$

The manufacturer's profit under the optimal production decision can be known through formulas (3), (4) and (1):

$$\Phi_A = [(m - nD + k(t) - \frac{s + G_B}{2})] \times \frac{m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2}}{2n} - \frac{rt}{2} \times \frac{m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2}}{2n}.$$

It can be known after finishing:

$$\Phi_A = \frac{(m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2})^2}{4n}. \tag{5}$$

Because of the existence of the big data platform, the information on the supply chain is completely symmetric. Therefore, the production decision of the supplier is based on the optimal production decision determined by the manufacturer. Then, the profit of the supplier under the manufacturer's optimal production decision can be obtained through formulas (3) and (4):

$$\Phi_B = \left(\frac{s + G_B}{2} - G_B\right) \times \frac{m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2}}{2n}.$$

It can be known after finishing:

$$\Phi_B = \frac{(s - G_B) \times (m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2})}{4n}. \tag{6}$$

According to the big data platform of the supply chain, in the case of decentralized decision-making, the manufacturer preferentially determines its own production decision and puts its own decision data on the big data platform. Based on the information obtained through the big data platform, the supplier produces the corresponding number of parts, and then the total profit of the supply chain at this time is:

$$\Phi = \Phi_A + \Phi_B = \frac{(m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2})^2}{4n} + \frac{(s - G_B) \times (m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2})}{4n}.$$

It can be known after finishing:

$$\Phi = \frac{(m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2})(m + k(t) - \frac{3s-G_B}{2} - \frac{rt}{2})}{4n}. \tag{7}$$

3.2 Common Decision-Making Profit Model

In the model of collaborative decision-making, suppliers and manufacturers rely on the big data platform for effective correlation, considering the overall interests of the supply chain to jointly determine the optimal production decision. Therefore, the function of the collaborative decision-making profit is expressed by the original profit function of the supplier and the manufacturer:

$$\varPhi' = \varPhi_A + \varPhi_B = (P_A - P_B) \times D - \frac{rt}{2} D + (P_B - G_B) \times D.$$

It can be known after finishing:

$$\varPhi' = \left(m - nD + k(t) - G_B - \frac{rt}{2}\right) \times D. \tag{8}$$

Find the partial derivative of D and make it equal to 0, it can be known that the optimal number of producers and suppliers at the time of collaborative decision-making is:

$$\begin{aligned} \frac{\partial \varPhi'}{\partial D} &= m - 2nD + k(t) - G_B - \frac{rt}{2} = 0, \\ D &= \frac{m + k(t) - G_B - \frac{rt}{2}}{2n}. \end{aligned} \tag{9}$$

The total profit model of supply chain can be known when the optimal production quantity has been determined under the condition of common decision through formula (8) and formula (9):

$$\varPhi' = \left(m - n \times \frac{m + k(t) - G_B - \frac{rt}{2}}{2n} + k(t) - G_B - \frac{rt}{2}\right) \times \frac{m + k(t) - G_B - \frac{rt}{2}}{2n}.$$

It can be known after finishing:

$$\varPhi' = \frac{\left(m + k(t) - G_B - \frac{rt}{2}\right)^2}{4n}. \tag{10}$$

4 Comparative Analysis

Through the above analysis, it obtains the profit model of the supplier and the manufacturer in the case of decentralized decision-making and collaborative decision-making when the manufacturer is fully committed to the big data cost.

In the case of decentralized decision-making, suppliers and manufacturers do not take into account the overall interests of the supply chain to maximize, but entirely from the perspective of their own interests to make decisions. Because of the existence of big data platform, although the suppliers and manufacturers make decisions discretely, the transmission of information flow is transparent. The manufacturer makes the decision first, and the supplier makes its own production decision according to manufacturer's production plan.

In the case of collaborative decision-making, to achieve the overall profit maximization of the supply chain is the premise of cooperation between suppliers and manufacturers, manufacturers take into account the supplier's production capacity constraints, logistics conditions and other objective factors, and then the two sides can achieve collaborative decision-making. The overall profit model of the supply chain is determined according to the market demand function of both parties, so that the production decision can be made with the maximum profit of the supply chain.

The profit model of the two cases is compared and analyzed, and then:

$$\begin{aligned}\Delta\Phi =&\Phi' - \Phi = \frac{(m + k(t) - G_B - \frac{rt}{2})^2}{4n} \\ &- \frac{(m + k(t) - \frac{s+G_B}{2} - \frac{rt}{2})(m + k(t) - \frac{3s-G_B}{2} - \frac{rt}{2})}{4n},\end{aligned}$$

for the convenience of calculation, let $e = m + k(t) - G - \frac{rt}{2}$, then there is

$$\begin{aligned}\Delta\Phi =& \frac{(e - G_B)^2}{4n} - \frac{(e - \frac{s+G_B}{2})(e - \frac{3s-G_B}{2})}{4n} \\ =& \frac{(e - G_B)^2}{4n} - \frac{(2e - s - G_B)(2e - 3s + G_B)}{16n} \\ =& \frac{(s - G_B)^2 + 4[(e - G_B)^2 - (e - s)^2]}{16n}.\end{aligned}$$

According to the above it can be known that s is the limit price that the market can afford, so there must be $G_B \langle s$, so, $(e - G_B)^2 - (e - s)^2 \rangle\, 0$, and easy to proof that $\Delta\Phi \rangle\, 0$, that is to say the profit in cooperative decision-making is higher than that in decentralized decision-making.

It can be known that manufacturers dominate the entire supply chain and completely control the supply chain's big data platform by analyzes and contrasts the two models, but it does not mean that the manufacturer can control the information flow entirely. Because in the case of decentralized decision-making, manufacturers do not take into account the conditions of suppliers when making decisions, so the interests of suppliers are damaged and lead to supply chain breaks. In the case of cooperative decision-making, manufacturers and suppliers, under the condition of ensure their own interests, determine the optimal number of production according to the conditions of both factors. After reaching the corresponding contract, the cooperation tends to be stable, which is to make the overall profit maximization of the supply chain as a prerequisite.

The above analysis can also prove that only cooperation can make the supply chain system to achieve the maximum profit, the joining of big data suppliers does not change the competition pattern of supply chain. Although manufacturers dominate the supply chain, choose to cooperate with suppliers will achieve a win-win situation.

5 Numerical Analysis

In this secondary supply chain, the overall profit of the supply chain with the raw material prices and market demand changes, Suppose that $K(t) = \frac{t^2}{2}$ and large data cost coefficient $r = 0.5$, consider the profit situation of supply chain when $t = 10$, raw material price and market limit price change, the concrete numerical analysis is shown in Table 1.

Table 1. Profitability of raw material price and market time price change

m	s	G_B	Φ	Φ'	$\Delta\Phi$	Rate of profit increase
35	25	10	812.5	1314.1	501.6	61.80%
35	25	15	820.3	1139.1	318.8	72.00%
35	30	10	664.1	1314.1	650	50.50%
35	30	15	675	1139.1	464.1	59.20%
40	25	10	962.5	1501.6	539.1	64.10%
40	25	15	970.3	1314.1	343.8	73.8%
40	30	10	801.6	1501.6	700	53.30%
40	30	15	812.5	1314.1	501.6	61.80%

It can be seen from the numerical analysis that when the price of raw materials is kept constant, the market limit price of the supplier is unchanged and the increase of the market limit price of the manufacturer will increase the overall profit of the two supply chains; The limit price of manufacture does not change and the increase of the market limit price of supplier has no effect on the overall profit of the supply chain. In addition, from the table can clearly see the common decision-making big data supply chain profits than traditional supply chain profits increase of at least 50.5%.

6 Conclusion

Big data theory has spawned many new service and industry models, and it is challenging for traditional industries to use big data to change their supply chains operating. How big data can be used to implement change, what should be targeted to gain competition advantage is the first issue enterprise should consider.

The model constructed in the paper focuses on the competition problem brought by big data suppliers join to the supply chain, and examines the supply chain cooperation mode from the profit perspective. It is not sharing big data cost by manufacturers and suppliers in the model, which is a common phenomenon in practice. In order to gain a dominant position in the supply chain, one enterprise chooses to bear the cost of big data alone to achieve the control of big data platform. But it does not means that the enterprise will be able to achieve an increase in profits, partly because of the cost of containment, on the other hand the stability of cooperation could be impacted after enterprise obtaining a dominant position, and the overall profit of the supply chain maybe reduced in that condition. Therefore, the supply chain member companies choose to cooperate with each other and make decisions collaboratively is the key to improve the overall profit of supply chain through introducing the third-party big data providers. The analysis of the article is also just to verify this conclusion.

Acknowledgements. This work was supported by the National Natural Science Foundation of China under Grant No. 61472027.

References

1. Chen DQ, Preston DS, Swink M (2015) How the use of big data analytics affects value creation in supply chain management. J Manage Inf Syst 32(4):4–39
2. Sanders NR (2016) How to use big data to drive your supply chain. Calif Manage Rev 58(3):26–48
3. Shi KR, He P, Xiao TJ (2011) A game approach to cooperative advertising in a two-stage supply chain. Ind Eng J 12:6–9
4. Tan KH, Zhan YZ et al (2015) Harvesting big data to enhance supply chain innovation capabilities: an analytic infrastructure based on deduction graph. Int J Prod Econ 165:223–233
5. Waller MA, Fawcett SE (2013) Click here for a data scientist: big data, predictive analytics, and theory development in the era of a maker movement supply chain. J Bus Logistics 34(4):249–252
6. Wang G, Gunasekaran A et al (2016) Big data analytics in logistics and supply chain management: certain investigations for research and applications. Int J Prod Econ 176:98–110
7. Wang HC, Cong JJ (2016) Study on coordination of cloud computation service supply chain under joint contracts. Logistics Technol 5:157–160
8. Wu C, Zhao DZ, Pan XY (2016) Comparison on dynamic cooperation strategies of a three-echelon supply chain involving big data service provider. Control Decis 7:1169–1171
9. Wu H, Han XH (2016) Production decisions in manufactures competing closed-loop supply chains under remanufacturing costs disruptions. Comput Integr Manuf Syst 4:1130–1140
10. Xin L, Jia Y (2011) Supply chain strategic partner selection based on game theory. Syst Eng 4:123–126
11. Xu L, Qi XU, Liu X (2015) A literature review on supply chain contracts under competing multi manufacturers. Technoeconomics Manage Res 7:3–7
12. Yang DJ, Bai Y (2015) Supply chain core firms competition and decentralization based on hotelling model. Proc Natl Acad Sci 102(4):1157–1162
13. Zhang A (2012) A comparative analysis of several demand function. J Jilin Province Inst Educ 10:153–154
14. Zhang TF, Zhang YM (2011) Analysis of knowledge sharing in supply chain node based on game theory. Stat Decis 17:184–186

Fuzzy GM(1,1) Model Based per Capital Income Predicted of Farmers in the World Natural and Cultural Heritage Area: Take Leshan City for an Example

Zhiyi Meng[1,2](✉), Caiyang Xu[3], and Hongxing Lan[1,2]

[1] Business College,Sichuan Agricultural University,
Dujiangyan 611800, People's Republic of China
zhiyimengscu@sina.com
[2] Southwest Poverty Reduction Center,
Sichuan Agricultural University, Chengdu 610065, People's Republic of China
[3] Business College, Sichuan Agricultural University,
Dujiangyan 611800, People's Republic of China

Abstract. The per-capita income of farmer is an important technology index which is the measured level of agriculture development and the status of farmers' life. This paper aims at developing a new fuzzy grey predicting model to predict the per-capita income of farmer. The fuzzy possibility is used to check the error of the proposed model, and the test results show that the accuracy of the model is quite high, providing a scientific basis for policy makers.

Keywords: The per-capita income of farmer · Fuzzy grey prediction model · Fuzzy possibility

1 Introduction

In the past few decades, China has experienced rapid economic growth and significant reduce of poverty [3]. However, at the same time, a lot of contradictions and problems have exposed. In recent years, the problem of slow growth in farmers' income has become the focus of the society from all walks of life. The difficulties faced by the current farmers' income growth have become a big obstacle to the virtuous cycle of the whole national economy. Therefore, the study of farmers' income growth is of far-reaching significance.

Since the reform and opening, China has carried out a series of significant reform in the countryside. In 1978, the rural reform which is market-orientation is a historic turning point of the agricultural development of China: not only break the bondage of the traditional system but also promoted the growth of farmers' income greatly. From 1978 to 2008, the per capita net income of farmers increased from 134 to 4761 [1]. In particular, over the past 30 years, the changes

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_83

of the growth of farmers' income in China are as follows. With the inflation factors deducted, from 1978 to 1985, it is growing at 15.2% yearly; from 1986 to 1991, at 2.7% yearly; from 1992 to 1996, at 5.6% yearly. Besides, after 1980, 1996 is the fastest growth year of farmers living income at the rate of 9% [7]. After 1997, enter the stage of slow growth while the growth rate of farmers' income was 4.6% which was the half of 1996. In 1998, their income continued going down while the speed of growth was only 4.3%. The year of 1999 and 2000 were the same and the speed of growth was 3.8% in 1999. The farmers' income was into a trough at 2.1% in 2000. In 2001, there was a recovery growth at 4.8%. However, it reduced to 4.3% in 2003. It achieved a steady increase from 2004 to 2008 and 2008 was at 8% [6].

In conclusion, there is a large change of growth range of per farming capita income and it will cause a negative influence to the development of rural economy even the whole national economy.

As a result, the prediction of per farming capita income becomes the focus of public concern. The key of long-term agriculture development is solving the problem of agriculture, rural areas and farmers (the three agriculture-related issues) reasonably and effectively. And the problem of per capita income of farmers is not only the core of this issue but also the key to solve the problem of the three agriculture-related issues. According to the history of the economic development of China, it's clear that the per capita income of farmers has an obvious change trend. In order to solve the problem of the three agriculture-related issues better, it needs to forecast the per capita income of farmers and grasp the change trend of it accurately as much as possible. However, some uncertain factors which are in the statistical process have become the chief problem of the prediction. For per capita income of farmers, there are many kinds of uncertain factors such as quantitative and non-quantitative factor, known and unknown factor.

Neglecting those non-quantified and unknown factors, the uncertain factors affecting the per capita income of farmers can be divided into two categories: (1) The part of production grew and sold by farmers themselves; (2) The production cost of agriculture.

There are many methods for forecasting the uncertain factors. In this paper, the prediction is based on the fuzzy gray prediction model.

The gray forecasting method has been applied in the modeling process of the dynamic system successfully in different fields such as agriculture, ecology, economy, statistics, industry and environment. In the absence of long-term historical data, it can use a system model to predict the incomplete or uncertain information.

According to the statistical method of per farming capita income from the statistical yearbook, the factors affecting per farming capita income (P) mainly include: the total income of farmers (TI), the total population of agriculture (IP). The relationship between them can be expressed by the following formula: $P = \frac{TI}{IP}$.

Leshan city of Sichuan province, as an area with World double-heritages, will be used as a case to study. The problem faced by the area with World double-heritages is how to exploit protectively. That is to say, the local natural and cultural heritages should be protected while the economy is developed and historical culture is promoted. Thus, there is a problem of degree. How to grasp this degree plays a vital role of the development of local economy and cultural protection which is also a model to the other areas with World double-heritages.

The purpose of this paper is to establish a fuzzy gray forecasting model to predict the per capita income of local farmers, get its changed trend and approve the effectiveness of the model.

2 Problem Statement

At present, per farming capita income is an important index which reflects the development of rural economy [5]. Therefore, it is very significant to predict and grasp accurately the changed trend of per farming capita income. This part will introduce the condition of per farming capita income of Leshan city and use it as an example to forecast per farming capita income with the fuzzy gray prediction model.

Due to the continuous application of the agricultural science and technology and the continuous improvement of the policies supporting agriculture by government, the total agricultural income of Leshan city has risen from 677.67 to 1992.15 million yuan in the past ten years. As shown in Fig. 1.

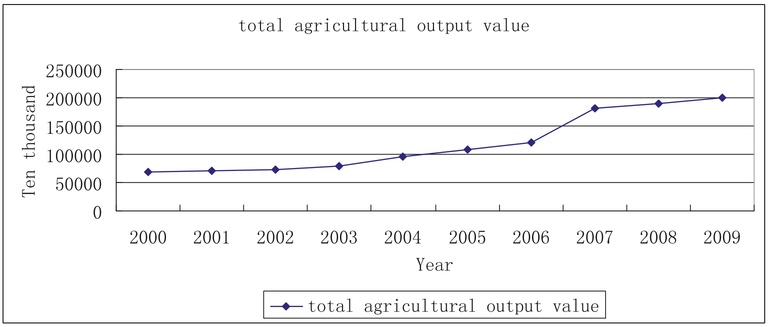


Fig. 1. The total agricultural output value of central district of Leshan city

From 2000 to 2009, the total agricultural output value of central district of Leshan city increased at a stretch. In addition to the year of 2006 and 2007, the data increased from 107.71 million yuan to 1.7 billion yuan. The overall increase in agricultural output value is in a small range and a lower growth rate. In order to a sustained and rapid growth of agricultural economy, the process of agricultural production should use advanced agricultural science and technology to increase the production and quality of agricultural products. At the same time, the policies published by government also need to be further improved.

Due to the implementation of the family control policy (a basic national policy) and the continuous improvement of the comprehensive qualities of local people, the agricultural population of central district had a tendency to reduce in Leshan city. As shown in Fig. 2.

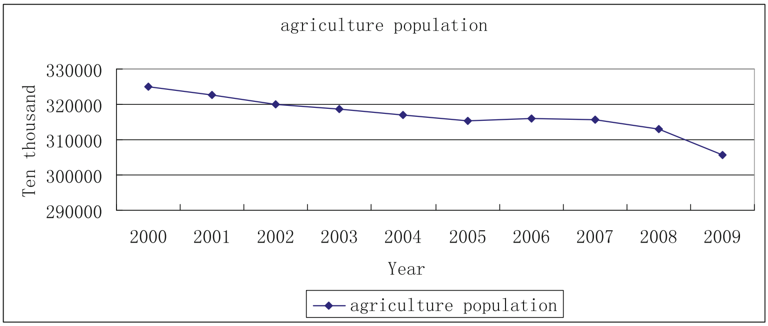


Fig. 2. The agriculture population of central district of Leshan city from 2000 to 2009

From 2000 to 2009, although there was a slight rebound from 2005 to 2007, the agricultural population of central district, Leshan city is of the gradual decline in the trend of change. The main reasons are as follows: On the one hand, affected by the Asian financial turmoil, from 2000 to 2005, crop prices were generally low and the enthusiasm of farmers was severely affected. However, since 2006, with the effect of financial crisis declining, the price of agricultural products rebounded and the enthusiasm of farmers was encouraged again. On the other hand, along with the rapid development of Chinese economy, the construction of urbanization is carried out in the vast rural areas. The acceleration of the urbanization process makes the local farming population decrease continuously. Thus, the agricultural population in central district, Leshan city also showed a decreasing trend. In summary, the total value of agricultural production of central district, Leshan city, has followed an increasing trend while the number of agricultural population has continued to decrease. Affected by the total value of agricultural production and the number of agricultural population, per capita income of farmers should also show an upward trend, as shown in Fig. 3.

From 2000 to 2009, the per capital income of farmers was increasing continuously with slow speed and low growth rate. The reasons are as follow: First, with the progress of science and technology, many kinds of advanced science and technology have been widely applied to agricultural production; second, the Chinese government has introduced various policies to support agricultural development which reduces the burden of farmers. For a sustained, rapid and steady growth of per capita income of farmers, it is necessary to improve the conversion rate of agricultural science and technology and perfect the various supporting policies of agriculture constantly.

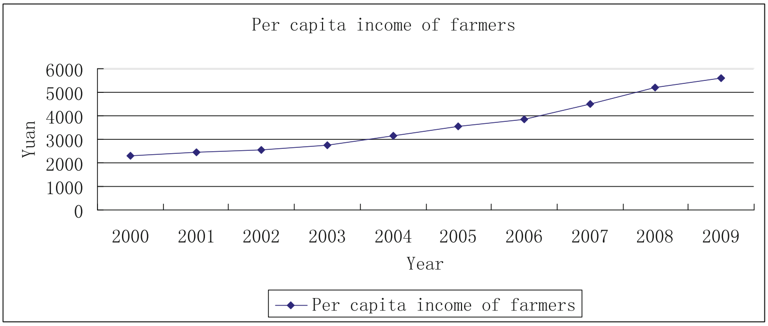


Fig. 3. The per capital income of farmers in central district, Leshan city between 2000 and 2009

3 Modeling

The gray system theory was first proposed by Deng [2], who used the differential equation as the forecasting model and the least square method to obtain the coefficient of the equation which has a wide range of application. In the gray forecasting system, the original data is defined as an accurate. However, in fact, the original data is not accurate which unable deal with the fuzzy phenomenon in reality effectively. In this paper, define the original data of per capita income of farmers as triangular fuzzy number and forecast the fuzzy numbers with gray forecasting system. Thus, this model consists of two parts: one is the fuzzy part; the other is the gray forecasting part.

3.1 Fuzzification

As mentioned earlier, the per capita income of farmers is affected by two factors: the total value of agricultural production (TI) and the agricultural population (IP). The relationship between them is: $P = \frac{TI}{IP}$.

When calculating the per capita income of farmers, this method has some shortcomings. Due to some indicators which unable get accurate enough data, it can only be estimated by experience. As a result, it will reduce the accuracy and scientificalness of the end result.

In the statistical process of the total income of farmers, some indicators unable be calculated accurately such as the part of self-production and marketing and the agricultural production cost of farmers. The self-production and marketing is the process to cost their own agricultural products. These agricultural products do not enter the market so that their value can not be measured accurately. In fact, the agricultural production cost includes labor costs while farmers own labor costs always unable be carried out. As a result, it unable be measured accurately. In these two parts of the statistical process, some indexes are often judged by experts' experience which is lack of precision.

Thus, the total income of the farmers in a region is only an estimate rather than an exact value. In this situation, it can be obscured by triangular fuzzy number to obtain a triangular fuzzy number of per capita incomes of farmers.

In order to solve the above problem, Zadeh proposed the fuzzy theory in 1965. Zadeh represent the fuzzy set by membership function following the idea that general set is represented by feature function [8].

Definition 1. Any mapping from domain X to closed interval [0, 1]. $\mu_A : X \to [0,1], x \to \mu_A : (X)$. Determine a fuzzy set A of X, μ_A is called the membership function of A, $\mu_A : (X)$ is called the membership degree of X to A, fuzzy set A is written as $A = \{(x, \mu_A) | x \in X\}$.

Definition 2. Set A as a fuzzy set in domain X. If $\forall_\alpha \in [0,1]$, the α- cut sets of A are all convex set so that the fuzzy set A is called as the convex fuzzy set.

Definition 3. If the fuzzy set M is a normal convex fuzzy set which is defined in the real field R, it is satisfied: (1) There is a unique point $x_0 \in$ R, making $\mu_M(x_0) = 1$(x_0 is called as the mean value of M); (2) $\mu_M(x)$ is continuous from left to right, then M is called as a fuzzy number. The meaning of the fuzzy number M is "the approximate real number of x_0".

From the definition of fuzzy number, the α- cut sets of M_α actually is a closed interval in real number field R: $M_\alpha = \{x \in \mathrm{R} | \mu_M(x) \geq \alpha\} = [m_\alpha^l, m_\alpha^r]$ m_α^l and m_α^r represent severally the left and right end points of the closed interval M_α.

The general expression of the fuzzy number M is $\mu_M(x) = \{L(x), l \leq x \leq m; R(x), m \leq x \leq r\}$; and $L(x)$ is the right continuous increasing function, $R(x)$ is the left continuous decreasing function, $0 \leq L(x), R(x) \leq 1$. If both the function $L(x)$ and $R(x)$ are linear functions, M is called as triangular fuzzy number, often denoted as $M(l, m, r)$. The total income of farmers is converted to triangular fuzzy number and expressed as $M(TI^-, TI, TI^+)$. The membership function shown in Fig. 4.

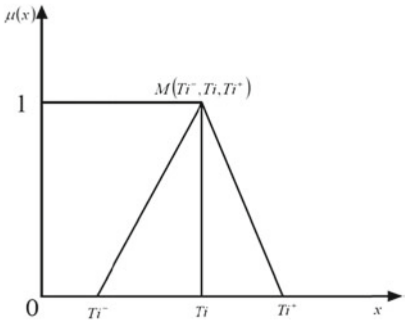


Fig. 4. The membership function of triangular fuzzy number

The relationship between the per capita income of farmers and the total income of farmers is:

$$P = \frac{TI}{TP}.$$

We can get the triangular fuzzy number of the historical data of the per capita income of farmers: $M(P^{-}, P, P^{+})$.

3.2 GM(1.1) Model

The GM(1,1) forecasting model is described below: Regard the original data of the per capita income of farmers as a set of original data series $p^{(0)}$, we can get:

$$p^{(0)} = \{(p^{(0)^-}(1), (p^{(0)}(1), p^{(0)^+}(1)), \cdots (p^{(0)^-}(n), (p^{(0)}(n), p^{(0)^+}(n))\}$$

and

$$p^{(1)} = \{(p^{(1)^-}(1), (p^{(1)}(1), p^{(1)^+}(1)), \cdots (p^{(1)^-}(n), (p^{(1)}(n), p^{(1)^+}(n))\}.$$

The relationship between the two sequences is as follows:

$$p^{(1)^-}(i) = \sum_{m=1}^{i} p^{(0)^-}(m), i = 1, 2, 3 \cdots, n,$$

$$p^{(1)}(i) = \sum_{m=1}^{i} p^{(0)}(m), i = 1, 2, 3 \cdots, n,$$

$$p^{(1)^+}(i) = \sum_{m=1}^{i} p^{(0)^+}(m), i = 1, 2, 3 \cdots, n,$$

$$\hat{p}^{(1)}(i+1) = (p^{0}(1) - \frac{y}{z})e^{-ai} + \frac{y}{z}.$$

According to the least squares method, we can conclude:

$$\hat{p}^{0}(1) = \hat{p}^{1}(1).$$

Then, the final formula is obtained:

$$\hat{p}^{0}(i+1) = \hat{p}^{1}(i+1) - \hat{p}^{1}(i).$$

4 Application

In order to get the triangular fuzzy number of per capita income of farmers in central district, Leshan city, it's necessary to fuzzify the historical data of per capita income of farmers in this area in the past ten years. The historical data are as shown in Table 1:

Table 1. The original data of per capita income of farmers in central district, Leshan city between 2000 to 2009

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Value (Yuan)	2306	2408	2573	2748	3154	3753	3829	4497	5192	5608

We suppose that:

$$\begin{aligned} P_0(k) &= (P_K^-, P_k, P_k^+), \\ P_k - P_k^- &= \alpha, \\ P_k^+ - P_k &= \beta. \end{aligned}$$

According to expert experience and related theories, suppose $\alpha = \beta = 100$. According to the above method, it can get the triangular fuzzy number of per capita income of peasants in central district, Leshan city by fuzzifing the data in the above table, as shown in Table 2:

Table 2. The triangular fuzzy number of original data of per capita income of peasants in central district, Leshan city between 2000 to 2009

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
P^-	2206	2308	2473	2648	3054	3653	3729	4397	5092	5508
P	2306	2408	2573	2748	3154	3753	3829	4497	5192	5608
P^+	2406	2508	2673	2848	3254	3853	3929	4497	5292	5708

By now, input the triangular fuzzy numbers of historical data of the per capita income of farmers in the central district of Leshan city to the gray forecasting model and get the predicted values are as Tables 3 and 4 shown:

Table 3. The predicted values of per capita income of farmers in the central district of Leshan city between 2001 to 2009

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009
P^-	2260	2532	2838	3180	3563	3993	4474	5014	5619
P	2166	2435	2737	3077	3459	3888	4371	4914	5524
P^+	2354	2630	2938	3283	3667	4097	4578	5114	5714

5 Result Analysis and Suggestion

Per capita income of farmers is an important indicator to reflect the development of agricultural economy and the living conditions of farmers. Whether the

Table 4. The predicted values of per capita income of farmers in the central district of Leshan city between 2001 to 2009

Year	2010	2011	2012	2013	2014	2015
P^-	6209	6980	7847	8821	9916	11147
P	6454	7302	8262	9349	10578	11968
P^+	6384	7132	7968	8902	9945	11111

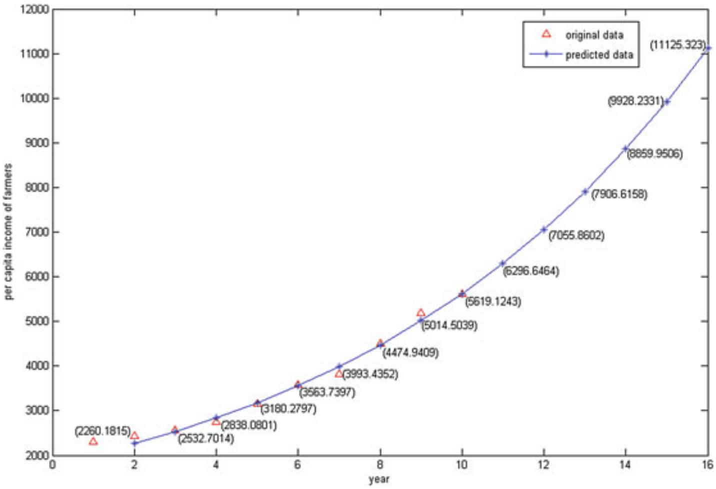


Fig. 5. The actual values and predicted values of P_K^-

per capita income of farmers has a sustained and rapid growth is important to the development of the entire national economy and the construction of a harmonious socialist society. According to the above forecast results, per capita income of farmers in central district, Leshan city is in the overall upward trend. In addition, the higher degree of fit between the predicted value and the original data indicates the validity of the fuzzy-gray forecasting model and also proves the scientificity of the predicted value. (As shown in Figs. 5, 6 and 7).

In the past decade, due to the continuous progress of agricultural science and technology and the continuous improvement of agricultural supporting policies, to a certain extent, it improved the productive rate of agriculture and reduced the burden on farmers. However, the per capita income of farmers is growing at a low speed with a low growth rate. Although there is a rising tendency, the per capita income of farmers is fluctuated because of some related economic factors.

In this view, the suggestions are as follows:

First of all, establish the work thinking of how to achieve stable increasing in income of farmers. The overall work thinking: the aim is to improve the income of farmers; the orientation is market; the support is resource; the basis is the structural adjustment; the driving force is science and technology; the way is

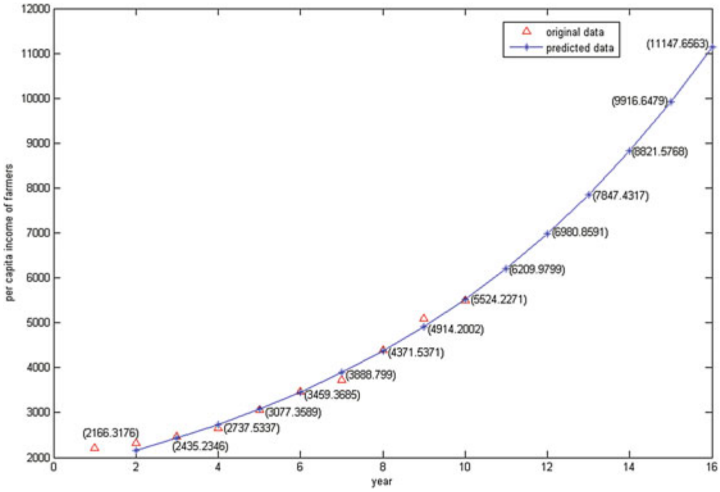


Fig. 6. The actual values and predicted values of P_K

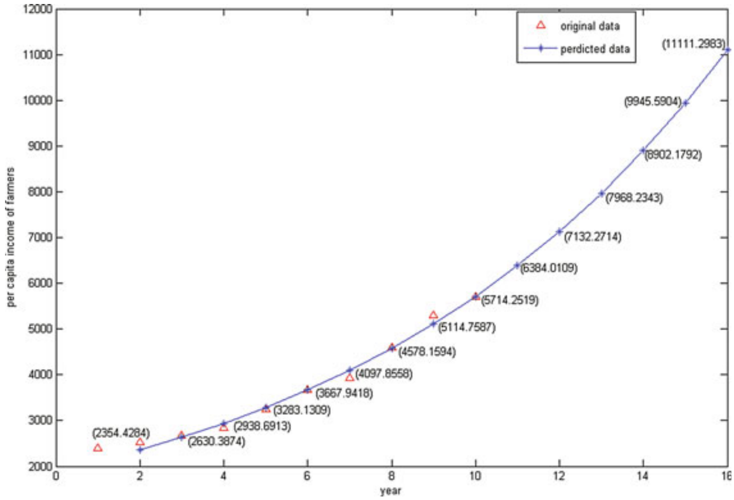


Fig. 7. The actual values and predicted values of P_K^+

industrialization and agricultural processing; the key is reducing agricultural population; develop green food, ecological agriculture and characteristic-brand agriculture vigorously [4].

Second, deepen the reform of the system and overcome the external factors that restrict farmers' income. China's history has proved repeatedly that when the state land and its property policy are correct and secure, when the enthusiasm of farmers is high, the income level has also a corresponding increase.

Third, the Government should implement the policy and measures to ensure that farmers' income increase. (1) It is of great significance for agricultural development and farmers' incomes to increase financial investment reasonably and carry out the agricultural subsidy policy. The funds for supporting agriculture should focus on poverty alleviation, agriculture foundational facilities, the research and application of agricultural technology and the green ecological agriculture. (2) Deepen the reform of rural financial. In the innovation of rural financial system, developing small and medium-sized banks and rural financial guaranty companies will solve the problems of rural finance fundamentally and provide a strong financial support for the development of three agriculture-related issues. It has an important role in regulation and protection to improve the production and income with the development of rural economy.

Finally, continue to promote the strategic adjustment of agricultural structure and improve the quality and efficiency of agriculture comprehensively. At this stage, improving the efficiency of agriculture should be market-oriented focusing on high quality and diversification of agricultural varieties. It means a change from yield-oriented to quality and efficiency.

6 Conclusion

Through the analysis of the per capita income of the farmers during the 30 years of reformation and opening, it is found that there are still problems in the growth process such as low growth rate and large fluctuation. Steady growth of farmers' income supports strongly for the economic with rapid growth of China and the smooth progress of reformation and opening.

In this paper, a forecasting model of the per capita income of farmers is established by fuzzy theory and gray forecasting system. As a case for our study, the central district, Leshan city provides with the relevant data. First, use the correlation theory of triangular fuzzy number to deal with the uncertain question; then use the gray forecasting model to predict the triangular fuzzy number which ensures the accuracy of the forecast results. At the same time, it provides a scientific basis for the formulation and implementation of relevant policies.

Acknowledgements. This work was supported by the Humanities and Social Sciences Foundation of the Ministry of Education (Grant No. 16YJC630089), and the Sichuan Center for System science and Enterprise Development Research (Grant No. Xq16C04).

References

1. Chen LQ (2013) Discussion on growth path about per capita annual net income of rural households of guangxi farmers. J South Agric 44(8):1402–1407
2. Deng J (1990) Application of grey system theory in China. In: International Symposium on Uncertainty Modeling and Analysis, 1990, Proceedings, pp 285–291
3. Lin JY (2012) Demystifying the Chinese Economy. Cambridge University Press, New York

4. Mao M, Chirwa EC (2006) Application of grey model GM(1, 1) to vehicle fatality risk estimation. Technol Forecast Soc Change 73(5):588–605
5. Tang HJ (2008) Models and strategic countes measures for circular agriculture development in China. J Agric Sci Technol 16:28–35
6. Xiang JH, Xiao L et al (2006) The situation and evaluation of China agriculture information website development. Agric Netw Inf 2:4–7
7. Yu X, Zhao G (2009) Chinese agricultural development in 30 years: a literature review. Front Econ China 4(4):633–648
8. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353

Eco-Cycle Comprehensive Operation Performance Evaluation–A Case Study of Baotou Steel Group

Yuyan Luo, Zhong Wang(✉), Yao Chen, Yahong Wang, and Jiming Xie

College of Management Science, Chengdu University of Technology,
Chengdu 610059, People's Republic of China
wangzhong2012@mail.cdut.edu.cn

Abstract. With the advent of circular economy, enterprises have no longer given priority to the economic benefits, but also focus on social responsibility and sustainable development. In order to accelerate the ecological change of enterprise system, every enterprise vigorously carries out green technology innovation and development. Therefore, how to evaluate the performance of ecological cycle has become a problem for management personnel. This paper takes Baotou steel group as an example, which is one of the largest steel companies in China, and establishes an ecological cycle performance evaluation index system from the aspects of economic operation, resource impact control and ecological circulation innovation. Based on that, the entropy-topsis model is used to evaluate the performance of ecological cycle in Baotou steel group. Through the evaluation, we find out that the ecological cycle performance of Baotou steel rose before 2010 and still at the middle and primary stage of its eco-cycle development. The results show the evaluation index system and method can help management personnel to realize the status of their company's eco-cycle development. It also can help policy makers to make better green policies and strategies, promoting the eco-cycle transformation in China.

Keywords: Eco-cycle · Operation performance evaluation · Entropy-topsis

1 Introduction

In recent years, with the transforming of the concept of ecological civilization construction, to realize low carbon development and cycle development has been the important goal of the international economic development, the development and the transformation of the target. The development direction of enterprise has changed from a single to rely on economic benefit as the guidance to the ecological cycle of performance oriented transformation. The evaluation of ecological cycle operating performance has become an important research topic, which

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_84

specifically reflected the enterprise transformation and put forward applicable suggestions for the future development of enterprise.

In the whole performance evaluation, building the evaluation index system is the critical part, reviewing the related literature, a large number of scholars have discussed the different content of the evaluation index system.

Such as the research using the three dynamic indicators including economic growth, environmental improvement and technical features to differentiate the enterprise ecological performance into 13 types including senior evolution to severe degradation [6]. Besides, there are some researches focusing on the circular economy developing standard of industrial park to evaluate, such as constructing the circular economy evaluation index system of industrial park including economic development, resource utilization, pollution control and material recycling indicators and using the TOPSIS method to evaluate [13] and based on AHP and FCE method, through economic development, the material reduction and circulation, pollution control and park management to build comprehensive evaluation model of circular economy of industrial parks [12], also, some researchers are committed to have assessment for the performance of enterprises and industrial parks, such as incorporating research and development (R&D) and marketing policy to build a innovative supply chain performance measurement system [15], to use discounted cash flow (DCF) and multi-attribute global inference of quality (MAGIQ) methods to evaluate the economic, environmental and social performances of eco-industrial park (EIP) pilot project [1], and based on principal component analysis and set pair analysis (SPA) and the fuzzy mathematics method to construct the performance evaluation system [5,8]. In addition, some use Delphi, LCA technique to select indicators to evaluate the green public procurement [14], setting the economic indicator and three generally applicable simplified environmental indicators (raw material consumption, energy consumption, and CO_2 emission) as the eco-efficiency indicators to evaluate the ecology efficiency of EIP [14], using the non-radial DEA model to evaluate the eco-efficiency [10], taking advantage of R cluster and factor analysis to have quantitative selection for indicators to build the green industry evaluation system including three criterion layers of green production, green consumption, and green environment [4], applying data envelopment analysis (DEA) to study the resource efficiency of cities [2], and presenting a compliance-based multidimensional trust evaluation system (CMTES) and also using improved TOPSIS technique to derive trustworthiness from compliance values [9].

Through the above review, a large number of methods such as the entropy-topsis, principal component analysis, gray correlation analysis, AHP and FCE, factor analysis, DEA model, the fuzzy comprehensive evaluation method and other evaluation methods have been widely used in the indicator system.

Based on the circular economy and industrial theory, this article builds the system of ecological cycle operation performance evaluation index and then takes the Baotou steel group as a case study, and analyzes the current eco-cycle operation performance and finally puts forward the corresponding countermeasures and suggestions.

2 Construction of Eco-Cycle Operation Performance System

The evaluation index system should accurately reflect the actual situation of enterprise's production and operation management. The establishment of index system is a key part of the whole ecological cycle operating performance's evaluation. It is essential to select the right index and guarantee the credibility of the results, therefore, this paper combines with the construction principles including system's scientificity and accuracy, effectiveness and comparability, stability and dynamic, operability, combination of qualitative and quantitative and so on, and based on a large number of research results of scholars and the actual work of ecological cycle transformation in enterprises, using analytic hierarchy process and expert-consulting method to select the indicators, builds the eco-cycle operational performance evaluation index system which includes economical operation, resource impact control, technology research and development and eco-cycle innovation. However, based on the performance evaluation of ecological recycling operation of Baotou steel, this research is aimed at finding that there exists missing data of the related index data in development of the technology and pollution control investment and resource compensation at resource impact control layer by collecting concrete data of Baotou steel such as annual report, social responsibility report and so on. Thus considering the actual operation of Baotou steel, we delete the indicators including R & D investment proportion of total amount, the number proportion of R & D in technology research and development level and pollution control investment growth rate and resource compensation fee increase rate in resource impact control layer. The final ecological recycling performance indicators are as follows in Fig. 1.

3 Entropy-Topsis Mode

On the basis of system theory, entropy is an ordered metric reflecting degree of disorder in the system [11]. Therefore, the smaller information entropy is, the lower disorder of information is, and the greater the effect of the information value and the weight information are, and vice versa [7]. Using entropy to reflect the size of information can come to the more objective and fair conclusions. Analytic hierarchy process (AHP) is usually used to determine the weighting factor of evaluation index in TOPSIS method, Because of its strong subjective factors, we use the improved entropy weight method to determine the weight of the evaluation index and to some extent avoid the influence of subjective factors. When the relative importance of evaluation index is difficult to determine, using the weight of entropy calculation index is more reasonable [3].

The general idea for using entropy weight method: To construct the index system and standard weighted matrix, and normalization of the relevant data into the model in the moment. Specific steps are as follows:

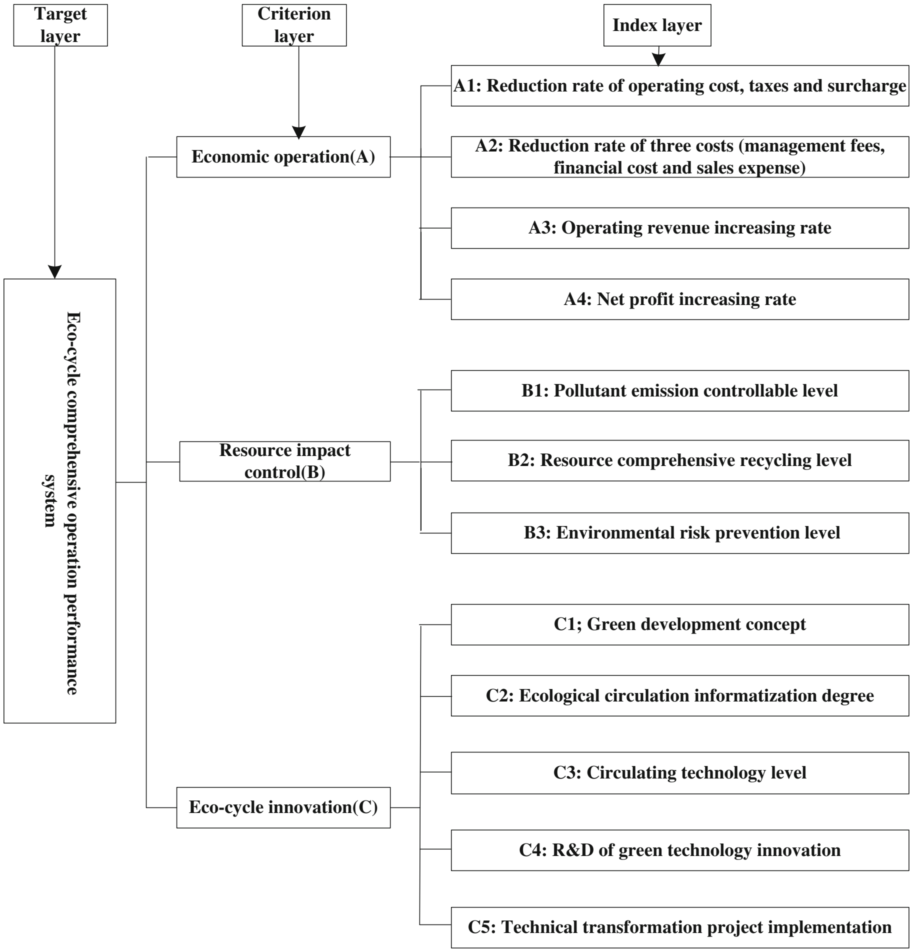


Fig. 1. Eco-cycle comprehensive operation performance system

(1) Data standardization
The evaluation system has n objects to be evaluated, m evaluation indicators, the evaluation matrix is $X = (x_{ab})_{n\times m}$ $(a = 1, 2, \cdots, n;\ b = 1, 2, \cdots, m)$.

For benefit type index:

$$y_{ab} = \frac{x_{ab} - \min x_b}{\max x_b - \min x_b}\ (1 \le a \le n, 1 \le b \le m). \tag{1}$$

For cost type index:

$$y_{ab} = \frac{\max x_b - x_{ab}}{\max x_b - \min x_b}\ (1 \le a \le n, 1 \le b \le m). \tag{2}$$

Normalization:

$$Q_{ab} = \frac{y_{ab}}{\sum_{a=1}^{n} y_{ab}} \; (1 \le a \le n, 1 \le b \le m). \tag{3}$$

(2) Calculation of entropy index

$$P_b = -1/\ln n \sum_{a=1}^{n} Q_{ab} \ln Q_{ab} \; (1 \le b \le m). \tag{4}$$

(3) Determine the index weight

$$\varphi_b = (1 - P_b)/\sum_{b=1}^{m} (1 - P_b), \tag{5}$$

where φ_b satisfies the two conditions $0 \le \varphi_b \le 1$ and $\sum_{b=1}^{m} \varphi_b = 1$:
(4) Standardized weighting matrix is obtained from normalized matrix and index weight.

$$W = (w_{ab})_{n \times m} = (\varphi_b y_{ab})_{n \times m}. \tag{6}$$

(5) Determine the positive and negative ideal solution
Positive ideal explanation imagines each index to achieve the most satisfactory solution, with W^+ and negative ideal solution is to imagine each index to achieve the most satisfactory solution, with W^-.

$$W^+ = \{\max w_{ab} | \, b \in V^+, \min w_{ab} | \, b \in V^-\} = \{w_1^+, w_2^+, \cdots, w_m^+\}, \tag{7}$$

$$W^- = \{\min w_{ab} | \, b \in V^+, \max w_{ab} | \, b \in V^-\} = \{w_1^-, w_2^-, \cdots, w_m^-\}, \tag{8}$$

where V^+= Benefit index set, V^-= Cost index set.
(6) Calculated distance
Each evaluation object to the government ideal sister distance with D_a^+, D_a^-, said, calculated by the following formula:

$$D_a{}^+ = \sqrt{\sum_{b=1}^{m} (w_{ab} - w_b^+)^2} (a = 1, 2, \cdots, n), \tag{9}$$

$$D_a{}^- = \sqrt{\sum_{b=1}^{m} (w_{ab} - w_b^-)^2} (a = 1, 2, \cdots, n). \tag{10}$$

(7) Calculate the relative closeness of each evaluation object

$$H_a = D_a{}^- / (D_a{}^- + D_a{}^+)(a = 1, 2, \cdots, n), \tag{11}$$

where H_a is the a object and how close the solution to, the value of H_a is between 0 and 1, the larger the value of H_a, the smaller the distance between the evaluation object and the positive ideal solution, the greater the distance from the negative ideal solution, the closer to the optimal level.

4 Case Study

Baotou steel group (hereinafter referred to as Baotou steel) is one of the earliest construction of iron and steel industrial bases after the founding of the People's Republic of China, it was first constructed in 1954 and was put into production in 1957. Turned into a corporate enterprise in 1998, it is the most important iron and steel industrial base and the largest rare earth industrial base in China, and one of the biggest rare earth scientific research and production bases as well. Always committed to diversify development, except to the major business of steel and rare earth, it also owns the mining and the steel industry. During the "twelfth five-year"? "big steel"? will be built and commit to become the world's largest production base of rare earth steel and the most competitive rare earth production and scientific research unit and its annual sales revenue will reach more than 100 billion yuan. By 2015, the group's total assets have amounted to 144.9 billion yuan, 33459 people have been employed and 22.501 billion yuan of sales income has been achieved. At present, Baotou steel has entered the ranks of tens of millions of tons of iron and steel enterprises in China, and has an important influence in the world. In recent ten years, Baotou steel has been committing to adhere to lead innovation, promoting the ecological cycle transformation and upgrading, keeping the environmental protection concept and setting the image of green development. Therefore, based on the practical situation of Baotou steel, evaluating its performance helps to future development selection and strategic decision-making.

(1) Data selection
According to the availability of data, this paper selects Baotou steel annual reports during the ten years from 2006 to 2015, part of the data are obtained by calculations, and others are obtained by the expert scoring method.

(2) Empirical results
In the process of analyzing of the index data, in order to eliminate the influence of dimension to the data, we use the formulas (1) and (2) to get the standardized data shown in Table 1, then take use of formula (3) to give normalized treatment to the data and finally get the weight P and the entropy ϕ.

As is shown in Table 2, the weight of the indicators for economic operation ranges between 0.02 and 0.033, the weight of the indicators for resource impact control is weighted between 0.06 and 0.12, the weight of the indicators for ecological circulation innovation range between 0.04 and 0.132. Therefore, the indicators of resource impact control have a great effect on the ecological cycle performance of Baotou steel.

Most of the indicators are at a low level, comprehensive index of ten years in a similar seven years posted schedule H below the 0.1 level. But it had a great reversal in 2010, whose comprehensive value of H reached the level of 0.04607, the value of H for economic operation even reached at 0.7 or so, and the value of H for resource impact control and the ecological circulation innovation were also at a high stage the ten years (Table 3).

Table 1. Standardized data

Year	A1	A2	A3	A4	B1	B2
2006	0.1107	0	0.0397	0.0911	0.1048	0.1347
2007	0.1159	0.1197	0.0007	0.0902	0.2048	0.1347
2008	0.1071	0.1112	0.0552	0.089	0.1476	0.1796
2009	0.1117	0.1118	0.0322	0.0866	0.1667	0.1856
2010	0	0.0913	0.7332	0.121	0.119	0.2156
2011	0.1103	0.1126	0	0.1401	0.1571	0.0749
2012	0.1109	0.119	0.0383	0.0949	0	0.0449
2013	0.1108	0.1041	0.0389	0.0905	0.0524	0.015
2014	0.1116	0.1147	0.0351	0.1967	0.0476	0.015
2015	0.111	0.1156	0.0268	0	0	0
Year	B3	C1	C2	C3	C4	C5
2006	0.0435	0.1019	0	0.1327	0.142	0.1393
2007	0	0.1019	0.1135	0.1467	0.1514	0.1333
2008	0	0	0.0054	0.1234	0.142	0.1274
2009	0.3913	0.051	0.0054	0.1292	0.0978	0.1298
2010	0.087	0.1911	0.1135	0.1467	0.142	0.1262
2011	0.0435	0.2229	0.2054	0.1257	0.1293	0.119
2012	0.0435	0.0127	0.1946	0.1118	0.142	0.1298
2013	0.1304	0.1274	0.1135	0.0012	0.041	0.0012
2014	0.1304	0.0637	0.1243	0	0.0126	0
2015	0.1304	0.1274	0.1243	0.0827	0	0.094

Table 2. Entropy weights of indicators

Item	Indicator	P	φ
Economical operation	A1	0.9542	0.0231
	A2	0.953	0.0236
	A3	0.4763	0.2634
	A4	0.9355	0.0325
Resource impact control	B1	0.8624	0.0692
	B2	0.8474	0.0768
	B3	0.7755	0.1129
Eco-cycle innovation	C1	0.765	0.1182
	C2	0.7384	0.1316
	C3	0.9004	0.0501
	C4	0.9	0.0503
	C5	0.9036	0.0485

Table 3. The closeness of the evaluation objects and the positive ideal solution

Year	Comprehensive operation		Economic operation		Resource impact control		Eco-cycle innovation	
	H	Rank	H	Rank	H	Rank	H	Rank
2006	0.0658	10	0.0417	6	0.0876	8	0.0839	8
2007	0.0773	5	0.018	10	0.1099	3	0.116	5
2008	0.0686	9	0.057	2	0.1077	4	0.0557	10
2009	0.1295	2	0.0361	7	0.3043	1	0.0596	9
2010	0.4607	1	0.7017	1	0.135	2	0.1497	2
2011	0.1093	3	0.0215	9	0.0854	9	0.1988	1
2012	0.0819	4	0.0418	5	0.039	10	0.1379	3
2013	0.0769	6	0.042	4	0.0979	5	0.1073	6
2014	0.071	8	0.044	3	0.0973	6	0.0904	7
2015	0.0765	7	0.0298	8	0.0936	7	0.1167	4

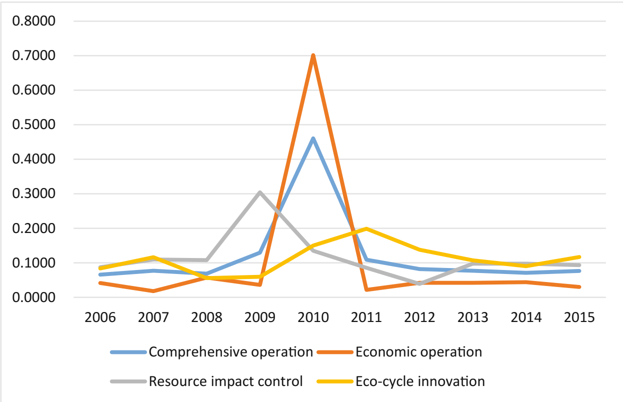


Fig. 2. The trend of closeness of each evaluation object and positive ideal solution

In order to reflect the current operating state of Baotou steel more intuitively and clearly, we make a brief line chart, it shown in Fig. 2.

The chart can reflect the development status of the last few years clearly and objectively, and combining with the specific content from annual report of Baotou steel, the following conclusions can be reached.

Baotou steel's comprehensive performance is at huge ups and downs. From 2006 to 2009, the overall ecological cycle operation of Baotou lagged behind slightly, but still grew steadily. But in the year of 2010, the business of Baotou steel had a great leap forward trend. The value of H rose from about 0.1 in 2009 to nearly 0.5. For Baotou steel, the year of 2010 is not only glorious, but also a sharp peak of development. The ecological cycle is good, which is the result of continuous accumulation. However, after 2010, Baotou steel business, abruptly suffered a great depression. In 2011, the value of H straightly dropped to about 0.05, and then was continuously declining during 2012, and was in a state between 0.05 and 0.1 from 2012 to 2015. So it's obvious to see that after Baotou steel's ecological cycle operates to a certain stage, the development of the operation will slow down, the challenge of industrial upgrading will increase.

As the same as the composite index, there also existed a parabola change trend for economic operation. The value of H reached 0.7 in 2010, but years after that it stayed between 0.01 and 0.057, which is similar to the analysis in (1). To be brief, Baotou steel business was in a condition of lower stage at first, but later went through such a huge revolution in 2010 and showed a trend of rapid increase. Thus it can be said that the year of 2010 is an important turning point in Baotou. Meanwhile, at the annual report. In previous years, Baotou steel has been committing to the construction of R & D center, and by the time of 2010, R & D center has been building more perfect. So it is not hard to explain before 2010, Baotou economy gradually increased, to show the reason for the substantial increase in 2010, and after a few years, Baotou research devoted to technology, may it lead to the decline in corporate profits.

Ecological circulation doesn't grow as such amazing speed as economic operation, but takes on a smooth curve of innovation. It presents a substantial increase in 2009 to 2011, makes a breakthrough in 2011, and gradually decline after 2011, finally is under the stable state. In general, compared with the 2000s, the ecological circulation innovation level in 2010s goes up. However, in comparison with the former two kinds of curve, the ecological circulation innovation reaches peak in 2011, but the economic performance and the comprehensive ecological cycle performance is relatively weak. Through the phenomenon and combining with the annual report of Baotou steel, during 2009 to 2011, Baotou steel has been promoting science and technology innovation and green development, introducing the resources recycling project to have an ecological cycle transformation, which leads to decline in the economic benefits. And then shows a trend of decline of ecological circulation innovation, it can draw that when ecological circulation innovation to achieve in a certain state, the growing rate slows down and breakthroughs are much more difficult to realize.

The trend of resource impact control is different from the former three. Through the line chart, we can clearly see that the sharp change of resource impact control is 2009 rather than 2010 or 2011, so, it appeared earlier than the year of vast change, and it can be seen in the line changes, the resource control increased in the previous year, then it would lead to eco-cycle comprehensive performance rising the next year, in the same way, when the resource impact control had a downward trend, after a year it would lead to a decline in the economic operation and eco-cycle comprehensive performance. Thus we can speculate that resource impact control is the foundation of ecological cycle of enterprise innovation and the daily basic implementation of enterprise resource recycling also determines the enterprise's operation and development.

Besides, usually, we can divide the H into five levels, lower stage ($N < 0.2$), primary stage ($0.2 \leq N < 0.4$), middle stage ($0.4 \leq N < 0.6$), relatively advanced stage ($0.6 \leq N < 0.8$), and advanced stage ($0.8 \leq N < 1.0$). So it's obvious to come out that the highest H values 0.46, the eco-cycle development of Baotou steel is in the middle and primary stage. Therefore, we can draw the conclusion that before 2010, the performance of Baotou steel's ecological cycling operation increased slowly, until 2010, the ecological cycle achieved a good level, but after 2010, industry innovation tended to be slow gradually, Baotou group operating performance gradually tumbled and the eco-cycle development of Baotou steel is in the middle and primary stage.

5 Conclusion

In view of the above empirical results, we conclude that the ecological performance of Baotou Steel rose in 2010, reached a peak in 2010, then decreased gradually, and the eco-cycle development of Baotou steel is in the middle and primary stage. And in the case of data support analysis, we confirm the reliability of the conclusion. So based on the conclusion of Baotou Steel Company, some suggestions are purposed as follows:

(1) Increase resource recycling and sustainable development. In the above analysis we can clearly observe the resource control impact indicator is the cornerstone of enterprise ecological cycle operation. Benign circulation of resources under the guidance of national policy is conducive to the development of enterprises. In the indexes we studied we can also know that controlling pollutant emissions, building chain impact prevention level, achieving effective material use and effective cycle are the measures to promote the development of enterprise resources. Of course, we can also through other ways such as optimization of industry leading and the construction of new production lines and other methods to improve the level of corporate resources impact control.

(2) Increase investment in green technology research and development. Drawn from the above analysis, eco-cycle innovation plays an important role in the eco-cycle comprehensive development of enterprises, so the ecological cycle innovation as a part of leading enterprise development is greatly important and cannot be ignored, therefore, increasing investment in green technology research and development is very important. The policy of the country now is to encourage the development of green industry and environmental protection enterprises, only to meet the mainstream trend is the real development, and Baotou as a high energy consumption and high pollution steel enterprise, designing and creating for the theme of environmental protection industry chain is the priority among priorities of Technology in recent years.

(3) Improve the energy structure. Increasing the recycling of resources and investment in green technology research and development are currently the most urgent task and measures based on the current. And for the future development of ten years, Baotou Steel Company should start in the long run, improve the energy structure of enterprise, deepen the reform and build

a green enterprise whose development conforms to the trend of the times. This is also the most useful measure in the future development.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Grant no. 71501019, 71501138, 71601164 and 71371130), Youth Program of Social Science Research of Sichuan Province for the Twelfth Five-year Plan (Grant no. SC15C005 and SC15C030), General Program of Education Department in Sichuan Province (Grant no. 16SB0071 and 16SB0049), General Program of Mineral Resources Research Center in Sichuan Province (Grant no. SCKCZY2014-YB04).

References

1. Chan FTS, Nayak A et al (2014) An innovative supply chain performance measurement system incorporating research and development (r & d) and marketing policy. Comput Ind Eng 69(1):64–70
2. Deilmann C, Lehmann I et al (2016) Data envelopment analysis of cities-investigation of the ecological and economic efficiency of cities using a benchmarking concept from production management. Ecol Ind 67:798–806
3. Dong J (2010) Analysis of external environment of transmission project based on improved entropy topsis method. Water Resour Power 3:152–154 (in Chinese)
4. Gongrong C, Yanghong D (2016) Performance evaluation of eco-industrial park based on the fuzzy mathematics method. J Hunan Univ Sci Technol 19:82–89
5. Jung S, Dodbiba G et al (2013) A novel approach for evaluating the performance of eco-industrial park pilot projects. J Cleaner Prod 39(5):50–59
6. Li P, Ren H, Zhao L (2005) Evaluation and analysis of enterprise eco-industrialization based on dynamic indicators. Quant Technica Econ 22(12):16–24
7. Mavi RK, Goh M, Mavi NK (2016) Supplier selection with shannon entropy and fuzzy topsis in the context of supply chain risk management. Procedia Soc Behav Sci 235:216–225
8. Rashidi K, Saen RF (2015) Measuring eco-efficiency based on green indicators and potentials in energy saving and undesirable output abatement. Energy Econ 50:18–26
9. Singh S, Sidhu J (2016) Compliance-based multi-dimensional trust evaluation system for determining trustworthiness of cloud service providers. Future Gener Comput Syst 67:109–132
10. Song X, Shen J (2015) The ecological performance of eco-industrial parks in shandong based on principal component analysis and set pair analysis. Resour Sci 37:546–554 (in Chinese)
11. Swendsen RH (2016) The definition of the thermodynamic entropy in statistical mechanics. Phys A Stat Mech Appl 467:67–73
12. Tarantini M, Loprieno AD, Porta PL (2011) A life cycle approach to green public procurement of building materials and elements: a case study on windows. Energy 36(5):2473–2482
13. Wu XQ, Wang Y et al (2008) Evaluation of circular economy development in industrial park based on eco-efficiency theory and topsis approach. Chin J Ecol 27(12):2203–2208
14. Ying Z, Hongzhi W, Guotai C (2016) Construction and application of green industry evaluation indicator system based on factor analysis. J Syst Manag 25:338–352
15. Yan L (2012) Assessment of circular economy in chlorine-alkali chemical industrial parks based on ahp-fce. Ind Technol Econ 31:151–155

Optimization of Closed-Loop Supply Chain Model with Part and Module Inventory Centers

YoungSu Yun[1(✉)], Anudari Chuluunsukh[1], Xing Chen[1], and Mitsuo Gen[2]

[1] Department of Business Administration, Chosun University, Gwangju, Korea
ysyun@chosun.ac.kr

[2] Fuzzy Logic Systems Institute, Tokyo University of Science, Katsushika, Japan

Abstract. In this paper, a closed-loop supply chain (CLSC) model with part and module inventory centers is proposed and optimized. Part supplier, module assembler, manufacturer, distribution center and retailer in forward logistics (FL) and customer, collection center, recovery center, secondary market and waste disposal center in reverse logistics (RL) are taken into consideration for constructing the proposed CLSC model. Especially, for the reuses of recovered part and module, part and module inventory centers are also considered. The proposed CLSC model is represented by a mathematical formulation, which is to minimize the sum of various costs resulting from each stage of FL and RL under satisfying various constraints. The mathematical formulation is implemented by an adaptive hybrid genetic algorithm (a-HGA) approach. In numerical experiment, various scales of the proposed CLSC model are presented and the performance of the a-HGA approach are compared with those of several conventional approaches. In conclusion, the efficiency of the proposed CLSC model and the a-HGA approach are proved.

Keywords: Closed-loop supply chain · Part and module inventory centers · Forward and reverse logistics · Adaptive hybrid genetic algorithm

1 Introduction

Closed-loop supply chain (CLSC) model is generally consisted of forward logistics (FL) and reverse logistics (RL). For FL, part supplier, manufacturer, distribution center and retailer are considered. For RL, customer, collection center, recovery (remanufacturing or refurbishing) center, secondary (used) market and waste disposal center are taken into consideration. In general, the objective for implementing the CLSC model can be divided into two aspects as follows:

- Constructing various facilities which can be used in FL and RL stages and considered in real-world situation.
- Optimizing the CLSC model.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_85

There are many conventional studies considering the above two aspects in the CLSC model [1–4,7,10,12].

For constructing various facilities, Fleischmann et al. [4] proposed the CLSC model with three activities of product production, resale and waste disposal in customer, reused market and disposer market, respectively. Amin and Zhang [1] proposed the CLSC model with reuse activity. For the reuse activity, they considered the new part from supplier as well as the reusable part from refurbishing center so that all parts are used for producing product at manufacturer in FL. Wang and Hsu [12] suggested the CLSC model with reuse activity, that is, recycler in RL classifies the returned product into reusable and unusable materials, respectively. The reusable materials are then reused at manufacturer in FL and unusable materials are disposed in landfill area. Similar to Wang and Hsu [12], Chen et al. [3] also suggested the CLSC model with various reuse activities. In this CLSC model, recycling center collects the returned product from customer and then classifies them into reusable and unusable products. The reusable product is reused at retailer in FL and the unusable product is disassembled into reusable and unusable materials. The reusable material is reused at manufacturer and the unusable material is treated at waste disposal plant in RL.

For optimizing the CLSC model, Amin and Zhang [1] and Wang and Hsu [12] minimized the total costs resulting from each stage of FL and RL. However, Chen et al. [3] maximized the total profit consisting of total revenue and total cost.

Although, the conventional studies mentioned above considered various activities of each facilities in FL and RL, they do not explained exactly how to use the reusable part (or material) for the facilities in FL. Therefore, in this paper, we propose a new CLSC model with two inventory centers (part and module inventory centers) so that the reusable part (or material) is exactly and effectively used for the facilities in FL. In Sects. 2 and 3, the proposed CLSC model is represented by a mathematical formulation, which is to minimize the sum of various costs resulting from each stage of FL and RL under satisfying various constraints. The mathematical formulation is implemented by an adaptive hybrid genetic algorithm (a-HGA) approach in Sect. 4. In Sect. 5 for numerical experiment, various scales of the proposed CLSC model are presented and the performance of the a-HGA approach are compared with those of several conventional approaches. Finally, in Sect. 6, as a conclusion, the efficiencies of the proposed CLSC model and the a-HGA approach are proved.

2 Proposed CLSC Model with Two Inventory Centers

The proposed CLSC model is consisted of part suppliers in areas 1, 2, 3 and 4, module assembler, product manufacturer, distribution center and retailer for FL and customer, collection center, recovery center, secondary market and waste disposal center in RL and two inventory centers (part and module inventory centers). Its conceptual structure is shown in Fig. 1.

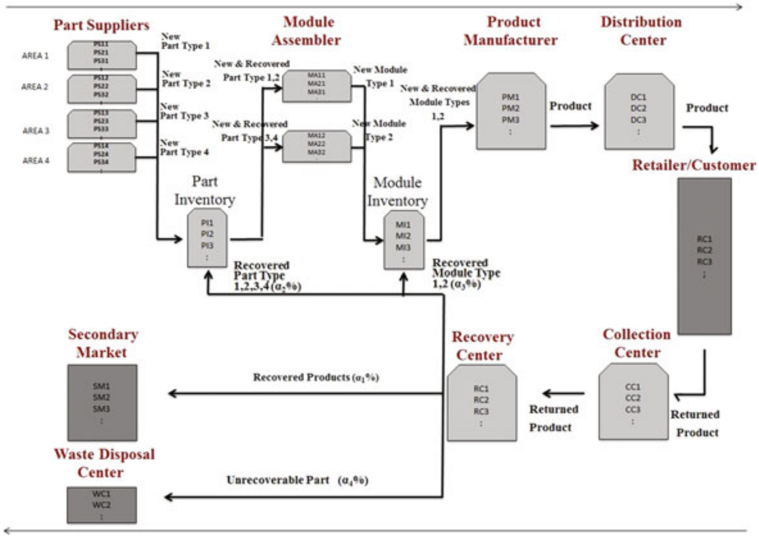


Fig. 1. Conceptual structure of the proposed CLSC model

The difference between the proposed CLSC model and the conventional CLSC models [1,3,12] is that the former considers two inventory centers and the latter does not taken into account them. The detailed logistics are as follows. Each part supplier at areas 1, 2, 3, and 4 respectively produces new part types 1, 2, 3, and 4 and then send them to part inventory center. Also, each module assembler assembles new module types 1 and 2 and then send them to module inventory center. Recovery center checks the returned product from collection center and then classifies them into recovered modules (recovered module types 1 and 2) and recovered parts (recovered part types 1, 2, 3, and 4). Each recovered part and module are sent to part and module inventory centers, respectively. Each inventory center has a function that new part and module in FL and recovered part and module in RL can be used for respectively assembling and producing module and product in module assembler and product manufacturer in FL. The recovered product and unrecovered module at recovery center are sent to secondary market and waste disposal center so that they are resold and land filled, respectively.

3 Mathematical Formulation

First, some assumptions are presented.

- Only single product is produced.
- The numbers of facility at each stage are already known.
- Among all facilities at each stage, only one facility should be opened.
- Fixed cost for operating each facility are different and already known.

- Unit handling cost considered at same stage is the same and already known.
- Unit transportation cost considered between each facility are different and already known.
- All products from customer are returned to collection center.
- The qualities of recovered part and module at recovery center are identical with those of new part and module.

Under considering the above assumption, the mathematical formulation for the CLSC model is proposed. Some indexes, parameters and decision variables are defined.

Index Set:
a : index of part supplier; $a \in A$;
p : index of area of part supplier; $p \in P$;
b : index of module assembler; $b \in B$;
m : index of area of module assembler; $m \in M$;
v : index of part inventory center; $v \in V$;
g : index of module inventory center; $g \in G$;
d : index of product manufacturer ; $d \in D$;
e : index of distribution center; $e \in E$;
y : index of retailer/customer; $y \in Y$;
j : index of collection center; $j \in J$;
z : index of recovery center; $z \in Z$;
i : index of secondary market; $i \in I$;
k : index of waste disposal center; $s \in S$.

Parameters:
BS_{ap} : fixed cost at part supplier a of area p;
BI_v : fixed cost at part inventory center v;
BM_{bm} : fixed cost at module assembler b of area m;
BV_g : fixed cost at module inventory center g;
BP_d : fixed cost at product manufacturer d;
BD_e : fixed cost at distribution center e;
BC_j : fixed cost at collection center j;
BR_z : fixed cost at recovery center z;
LS_{ap} : unit handling cost at part supplier a of area p;
LI_v : unit handling cost at part inventory center v;
LM_{bm} : unit handling cost at module assembler b of area m;
LV_g : unit handling cost at module inventory center g;
LP_d : unit handling cost at product manufacturer d;
LD_e : unit handling cost at distribution center e;
LC_j : unit handling cost at collection center j;
LR_z : unit handling cost at recovery center z;
TSI_{apv} : unit transportation cost from part supplier a of area p to part inventory center v;
TIM_{vbm} : unit transportation cost from part inventory center v to module assembler b of area m;

TMV_{bm} : unit transportation cost from module assembler b of area m to module inventory center g;
TVP_{gd} : unit transportation cost from module inventory center g to product manufacturer d;
TPD_{de} : unit transportation cost from product manufacturer d to distribution center e;
TDT_{ey} : unit transportation cost from distribution center e to retailer/customer y;
TTC_{yj} : unit transportation cost from retailer/customer y to collection center j;
TCV_{jz} : unit transportation cost from collection center j to recovery center z;
TVU_{zk} : unit transportation cost from recovery center z to secondary market k;
TVI_{zv} : unit transportation cost from recovery center z to part inventory center v;
TVV_{zg} : unit transportation cost from recovery center z to module inventory center g;
TUW_{zk} : unit transportation cost from recovery center z to waste disposal center k;

Decision Variables:
ii_v : handling capacity at part inventory center v;
im_{bm} : handling capacity at module assembler o of area e;
iv_g : handling capacity at module inventory center g;
ip_d : handling capacity at product manufacturer d;
id_e : handling capacity at distribution center e;
ir_y : handling capacity at retailer/customer y;
ic_j : handling capacity at collection center j;
iv_z : handling capacity at recover center z;
iu_i : handling capacity at secondary market i;
iw_k : handling capacity at waste disposal center k

$$ps_{ap} = \begin{cases} 1, \text{ if part supplier } p \text{ of area } a \text{ is opened} \\ 0, \text{ otherwise} \end{cases}$$
$$pi_v = \begin{cases} 1, \text{ if part inventory center } v \text{ is opened} \\ 0, \text{ otherwise;} \end{cases}$$
$$pm_{bm} = \begin{cases} 1, \text{ if module assembler } b \text{ of area } m \text{ is opened} \\ 0, \text{ otherwise;} \end{cases}$$
$$pv_g = \begin{cases} 1, \text{ if module inventory center } g \text{ is opened} \\ 0, \text{ otherwise;} \end{cases}$$
$$pp_d = \begin{cases} 1, \text{ if product manufacturer } d \text{ is opened} \\ 0, \text{ otherwise;} \end{cases}$$
$$pd_e = \begin{cases} 1, \text{ if distribution center } e \text{ is opened} \\ 0, \text{ otherwise;} \end{cases}$$
$$pc_j = \begin{cases} 1, \text{ if distribution center } j \text{ is opened} \\ 0, \text{ otherwise;} \end{cases}$$

$$pr_z = \begin{cases} 1, \text{ if distribution center } z \text{ is opened} \\ 0, \text{ otherwise.} \end{cases}$$

Objective function and constraints are as follows:
Minimize Total Cost (TC)

$$\begin{aligned} TC = &\sum_p \sum_a (BS_{ap} ps_{ap}) + \sum_v (BI_v \times pi_v) + \sum_b \sum_m (BM_{bm} \times pm_{bm}), \\ &+ \sum_g (Bv_g \times pv_g) + \sum_d (BP_d \times pp_d) + \sum_e (BD_e \times pd_e) + \sum_j (BC_j \times pc_j) \\ &+ \sum_z (BR_z \times pr_z) + \sum_a \sum_p (LS_{pa} \times is_{pa} \times ps_{pa}) + \sum_v (LI_v \times ii_v \times pi_v) \\ &+ \sum_m \sum_b (LM_{bm} \times im_{bm} \times pm_{bm}) + \sum_g (LV_v \times iv_g \times pv_g) + \sum_d (LP_d \times ip_d \times pp_d) \\ &+ \sum_e (LD_e \times id_e \times pd_e) + \sum_j (LC_j \times ic_j \times pc_j) + \sum_z (LR_z \times ir_z \times pr_z) \qquad (1) \\ &+ \sum_p \sum_a \sum_v (\mathrm{TSI}_{apv} \times is_{pa} \times ps_{pa} \times pi_v) + \sum_v \sum_m \sum_b (\mathrm{TIM}_{vbm} \times ii_v \times pi_v \times pm_{bm}) \\ &+ \sum_m \sum_b \sum_g (\mathrm{TMV}_{eoc} \times im_{bm} \times pm_{bm} \times pv_g) + \sum_g \sum_d (\mathrm{TVP}_{gd} \times iv_g \times pv_g \times pd_d) \\ &+ \sum_d \sum_e (\mathrm{TPD}_{de} \times p_d \times pp_d \times pd_e) + \sum_e \sum_y (\mathrm{TDT}_{ey} \times id_e \times pd_e) \\ &+ \sum_y \sum_j (\mathrm{TTC}_j \times ic_j \times pc_j) + \sum_j \sum_z (\mathrm{TCV}_{jz} \times ir_z \times pr_z \times pv_j) \\ &+ \sum_z \sum_v (\mathrm{TVI}_{zv} \times ir_z \times pr_z \times pi_v) + \sum_z \sum_g (\mathrm{TVV}_{zg} \times ir_z \times pr_z \times pv_g) \\ &+ \sum_z \sum_i (\mathrm{TVU}_{zi} \times ir_z \times pr_z) + \sum_z \sum_k (\mathrm{TVW}_{zk} \times ir_z \times pr_z). \end{aligned}$$

$$\text{s. t.} \sum_{\mathrm{p}} \mathrm{ps}_{\mathrm{ap}} = \mathbf{1}, \forall \mathbf{p} \in \mathbf{P}, \tag{2}$$

$$\sum_{\mathrm{v}} \mathrm{pi}_{\mathrm{v}}, \tag{3}$$

$$\sum_{\mathrm{m}} \mathrm{pm}_{\mathrm{bm}} = \mathbf{1}, \forall \mathbf{m} \in B, \tag{4}$$

$$\sum_{\mathrm{g}} \mathrm{pv}_{\mathrm{g}} = \mathbf{1}, \tag{5}$$

$$\sum_{\mathrm{d}} \mathrm{pp}_{\mathrm{d}} = \mathbf{1}, \tag{6}$$

$$\sum_{\mathrm{e}} \mathrm{pd}_{\mathrm{e}} = \mathbf{1}, \tag{7}$$

$$\sum_{\mathrm{j}} \mathrm{pc}_{\mathrm{j}} = \mathbf{1}, \tag{8}$$

$$\sum_{\mathrm{z}} \mathrm{r}_{\mathrm{z}} = \mathbf{1}, \tag{9}$$

$$\sum_{\mathrm{ap}} (\mathrm{is}_{\mathrm{ap}} \times \mathrm{ps}_{\mathrm{ap}}) - \sum_{\mathrm{v}} (\mathrm{ii}_{\mathrm{v}} \times \mathrm{pi}_{\mathrm{v}}) \leq 0 \quad \forall p \in P, \tag{10}$$

$$\sum_{\mathrm{v}} (\mathrm{ii}_{\mathrm{v}} \times \mathrm{pi}_{\mathrm{v}}) - \sum_{\mathrm{bm}} (\mathrm{im}_{\mathrm{bm}} \times \mathrm{pm}_{\mathrm{bm}}) \leq 0 \quad \forall b \in B, \tag{11}$$

$$\sum_{\mathrm{bm}} (\mathrm{im}_{\mathrm{bm}} \times \mathrm{pm}_{\mathrm{bm}}) - \sum_{\mathrm{g}} (\mathrm{iv}_{\mathrm{g}} \times \mathrm{pv}_{\mathrm{g}}) \leq 0 \quad \forall b \in B, \tag{12}$$

$$\sum_{\mathrm{g}} (\mathrm{iv}_{\mathrm{g}} \times \mathrm{pv}_{\mathrm{g}}) - \sum_{\mathrm{d}} (\mathrm{ip}_{\mathrm{d}} \times \mathrm{pp}_{\mathrm{d}}) = 0, \tag{13}$$

$$\sum_{\mathrm{d}} (\mathrm{ip}_{\mathrm{d}} \times \mathrm{pp}_{\mathrm{d}}) - \sum_{\mathrm{e}} (\mathrm{id}_{\mathrm{e}} \times \mathrm{pd}_{\mathrm{e}}) = 0, \tag{14}$$

$$\sum_{\mathrm{e}} (\mathrm{id}_{\mathrm{e}} \times \mathrm{pd}_{\mathrm{e}}) - \sum_{\mathrm{y}} \mathrm{qr}_{\mathrm{y}} = 0, \tag{15}$$

$$\sum_{\mathrm{y}} \mathrm{qr}_{\mathrm{y}} - \sum_{\mathrm{j}} (\mathrm{ic}_{\mathrm{j}} \times \mathrm{pc}_{\mathrm{j}}) = 0, \tag{16}$$

$$\sum_{\mathrm{j}} (\mathrm{ic}_{\mathrm{j}} \times \mathrm{pc}_{\mathrm{j}}) - \sum_{\mathrm{z}} (\mathrm{ir}_{\mathrm{z}} \times \mathrm{pr}_{\mathrm{z}}) \geq 0, \tag{17}$$

$$\sum_{\mathrm{v}} (\mathrm{ii}_{\mathrm{v}} \times \mathrm{pi}_{\mathrm{v}}) - a_1\% \sum_{\mathrm{z}} (\mathrm{ir}_{\mathrm{z}} \times \mathrm{pr}_{\mathrm{z}}) \geq 0, \tag{18}$$

$$\sum_{\mathrm{g}} (\mathrm{iv}_{\mathrm{g}} \times \mathrm{pv}_{\mathrm{g}}) - a_2\% \sum_{\mathrm{z}} (\mathrm{ir}_{\mathrm{z}} \times \mathrm{pr}_{\mathrm{z}}) \geq 0, \tag{19}$$

$$\sum_{\mathrm{i}} \mathrm{pu}_{\mathrm{i}} - a_3\% \sum_{\mathrm{z}} (\mathrm{ir}_{\mathrm{z}} \times \mathrm{pr}_{\mathrm{z}}) \geq 0, \tag{20}$$

$$\sum_{\mathrm{k}} \mathrm{pw}_{\mathrm{k}} - a_4\% \sum_{\mathrm{z}} (\mathrm{ir}_{\mathrm{z}} \times \mathrm{pr}_{\mathrm{z}}) \geq 0, \tag{21}$$

$$\mathrm{ps}_{\mathrm{ap}} = \{0, 1\}, \quad a \in A, \quad \forall p \in P, \tag{22}$$

$$\mathrm{pi}_{\mathrm{v}} = \{0, 1\}, \quad \forall i \in I, \tag{23}$$

$$\mathrm{pm}_{\mathrm{bm}} = \{0, 1\}, \quad \forall n \in M, \quad b \in B, \tag{24}$$

$$\mathrm{pv}_{\mathrm{g}} = \{0, 1\}, \quad \forall g \in G, \tag{25}$$

$$\mathrm{pp}_{\mathrm{d}} = \{0, 1\}, \forall d \in D, \tag{26}$$

$$\mathrm{pd}_{\mathrm{e}} = \{0, 1\}, \quad \forall e \in E, \tag{27}$$

$$\mathrm{pc}_{\mathrm{j}} = \{0, 1\}, \quad \forall j \in J, \tag{28}$$

$$\mathrm{pr}_{\mathrm{z}} = \{0, 1\}, \quad \forall z \in Z, \tag{29}$$

$$\begin{aligned} &\mathrm{ps}_{\mathrm{ap}}, \mathrm{pi}_{\mathrm{v}}, \mathrm{pm}_{\mathrm{bm}}, \mathrm{pp}_{\mathrm{d}}, \mathrm{pd}_{\mathrm{e}}, \mathrm{pc}_{\mathrm{y}}, \mathrm{pr}_{\mathrm{z}} \mathrm{pu}_{\mathrm{i}}, \mathrm{pw}_{\mathrm{k}} \geq 0, \forall p \in P, \quad \forall a \in A, \quad \forall v \in V, \\ &\forall m \in M, \quad \forall b \in B, \quad \forall d \in D, \quad \forall e \in E, \quad \forall y \in Y, \quad \forall z \in Z, \quad \forall i \in I, \quad \forall k \in K, \\ &\forall j \in J. \end{aligned} \tag{30}$$

The objective function of Eq. (1) is to minimize the total sum of fixed costs, handling costs and transportation costs. Equations (2)–(9) means that only one facility is opened at each stage. Equation (10) implies that the sum of the handling capacity at each supplier in areas 1, 2, 3 and 4 is the same or greater than that of the module inventory center. The same meaning is considered in Eqs. (11)–(17). Equation (18) implies that the sum of the recovered products at each recovery center is the same or greater than that of the recoverable products with $a_1\%$ at each collection center. Equation (19) restricts that the sum of the handling capacity at all part inventory center is the same or greater than that of the recoverable product with $a_2\%$ at all recovery centers. Equations (20)–(21) indicate the same meanings of the Eqs. (18) and (19). Equations (22)–(29) restrict the variables to integers 0 and 1. Equation (30) means non-negativity.

4 A-HGA Approach

The mathematical formulation is implemented using the a-HGA approach. The a-HGA approach is a hybrid approach with adaptive scheme. For the hybrid approach, conventional GA and Cuckoo search (CS) are used and for the adaptive scheme, Srinivas and Patnaik's approach [11] is used. Using the hybrid approach can achieve a better improvement of solution rather than using a single approach does. By using the adaptive scheme, the rates of crossover and mutation operators used in GA are automatically regulated. The detailed implementation procedure [5,6,8] is as follows.

Step 1. GA approach
- **Step 1.1.** Representation 0–1 bit representation scheme is used for effectively representing opening/closing decision of all facilities at each stage.
- **Step 1.2.** Selection Elitist selection strategy in enlarged sampling space is used
- **Step 1.3.** Crossover Two-point crossover operator ($2X$) is used.
- **Step 1.4.** Mutation Random mutation operator is used.
- **Step 1.5.** Reproduce offspring

Step 2. CS approach Apply Levy flight scheme [8] to the offspring of GA and produce new solution.

Step 3. Adaptive scheme Apply the adaptive scheme used in Srinivas and Patnaik [11] to regulate crossover and mutation rates.

Step 4. Termination condition If pre-determined stop condition is satisfied, then stop all Steps, otherwise go to Step. 1.2.

5 Numerical Experiments

In numerical experiment, three scales of the proposed CLSC model are presented. Each scale has various sizes of part suppliers in areas 1, 2, 3 and 4, par inventory center, module assembler in area 1 and 2, module inventory center, product manufacturer, distribution center and retailer in FL and customer, collection center, recovery center, secondary market, and waste disposal center in RL. The detailed sizes of each scale are showed in Table 1. For each scale, 1,500 products are produced in FL and handled in RL. The rates at recovery center for handling the returned products from collection center are as follows: $\alpha_1 = 60\%$,

Table 1. Various scales of the proposed CLSC model

Scale	Part supplier				Part inventory center	Module assembler		Module inventory center	Product manufacturer	Distribution center	Retalier/customer	Collection center	Recovery center	Secondary market	Waste disposal center
	1	2	3	4		1	2								
1	4	4	4	4	1	3	3	1	4	3	10	4	3	10	1
2	12	12	12	12	2	8	8	2	12	8	15	12	8	15	2
3	23	23	23	23	3	20	20	3	23	20	25	20	20	25	3

Table 2. Each approach for comparison

Approach	Description
GA	Conventional GA [5]
HGA	Conventional HGA by Kanagaraj et al. [6]
a-HGA	Proposed approach in this paper
Lingo	Conventional optimization solver by Lindo Systems [8]

Table 3. Measures of performance comparison

Measure	Description
Best solution	Best value of the objective functions under satisfying all constraints
Average solution	Averaged values of the objective functions under satisfying all constraints
Average time	Average value of the CPU time (Sec.) used for running each approach
Percentage difference	The difference of the best solutions of GA, HGA and a-HGA when compared with that of Lingo

$\alpha_2 = 20\%$, $\alpha_3 = 15\%$ and $\alpha_4 = 5\%$, for recoverable products, parts, modules and unrecoverable modules, respectively.

To prove the efficiency of the a-HGA approach, two conventional approaches (GA, HGA) and Lingo [9] as a benchmark are used and they are summarized in Table 2. The performances of each approach are compared using various measures of performance as shown in Table 3.

All approaches ran on a same computation environment (IBM compatible PC 1.3 Ghz processor-Intel core $I5 - 1600$ CPU, 4 GB RAM, OS-X EI) and programmed by MATLAB version R2014a. The parameter settings for GA, HGA and a-HGA are as follows: For each scale, total number of generations is 1,000, the population size is 20, the crossover and mutation rates are 0.5 and 0.3 respectively. Total 30 trials were independently done for eliminating the randomness in

Table 4. Computation results of each approach

		Best solution	Average iteration	Average solution	Average time	Percentage diff.
Scale 1	GA	149,789	15	150,754	8.1	−7.41%
	HGA	149,789	17	150,520	8.30	−7.41%
	a-HGA	149,789	21	149,789	138.50	−7.41%
	LINGO	161,777	-	-	-	0.00%
Scale 2	GA	149,868	35	152,381	8.33	−10.12%
	HGA	149,852	42	152,679	8.83	−10.13%
	a-HGA	149,765	54	151,616	193.70	−10.18%
	LINGO	166,747	-	-	-	0.00%
Scale 3	GA	151,572	53	153,919	9.00	−9.31%
	HGA	149,943	54	153,610	9.07	−10.27%
	a-HGA	149,895	42	153,365	313.30	−10.30%
	LINGO	167,103	-	-	-	0.00%

the search processes of the GA, HGA and a-HGA. Table 4 shows the computation results by GA, HGA, a-HGA and Lingo.

In the scale 1 of Table 4, the a-HGA including the GA and HGA has the same re-sult and their performances are greater than that of Lingo in terms of the best solution and percentage difference. In terms of the average solution, the a-HGA shows the best performance compared with the GA and HGA. However, in terms of the average time, the a-HGA shows the worst performance and the GA is the best performer. In scale 2, the performance of the a-HGA is more efficient than the GA and HGA in terms of the best solution and average solution. The difference in terms of the percentage difference means that the a-HGA is 0.05% and 0.06% advantageous compared with the HGA and GA. However, in terms of the average time, the a-HGA is the slowest and the GA is the quickest.

Similar results are also shown in the scale 3, that is, the a-HGA shows the best performances in terms of the best solution, average solution and percentage difference when compared with the GA, HGA and Lingo. However, the search speed of the a-HGA is about thirty times slower than those of the GA and HGA.

Figure 2 shows the convergence behaviors of GA, HGA and a-HGA until the generation number is reached to 200.

In Fig. 2, all approaches show rapid and various convergence behaviors during initial generations. However, after about 50 generations, each approach does not show any more convergence behaviors and the a-HGA shows to be more efficient behaviors than the GA and HGA.

The result of the detailed material flows in the a-HGA for the scale 3 is shown in Fig. 3. The opened facilities at each stage are displayed as white-coloured boxes. The 975 new parts in each area are produced and then sent to part inventory center. The recovery center recovers the quality of the returned products

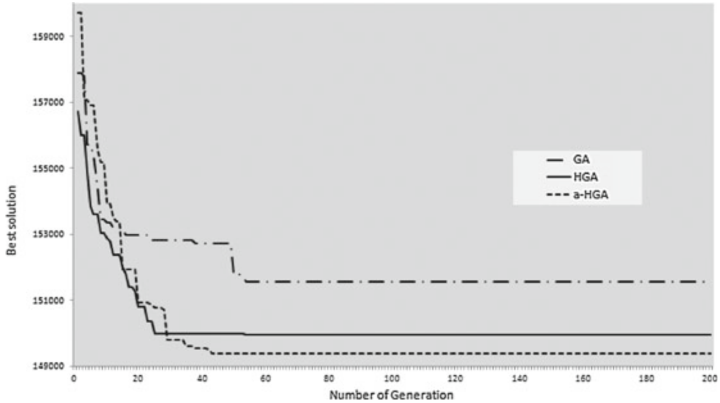


Fig. 2. Convergence behaviors of GA, HGA and a-HGA

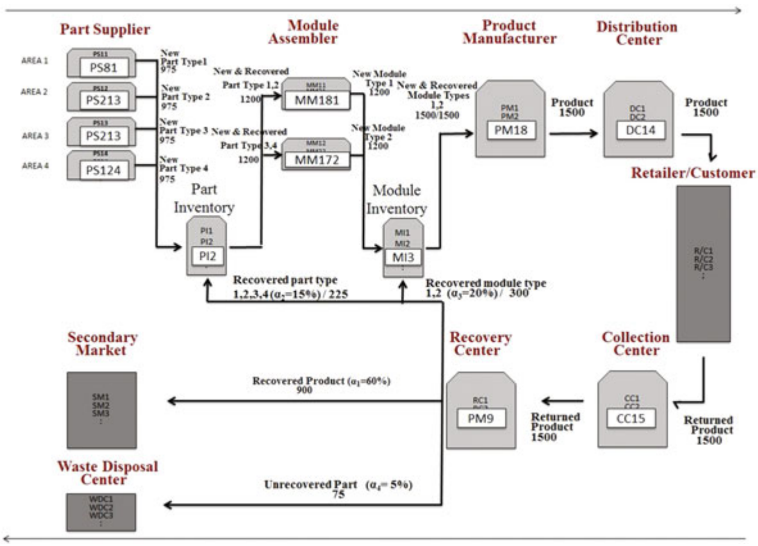


Fig. 3. Detailed material flows and facility numbers opened at each stage

and then sent 225 recovered parts to part inventory center. Part inventory center stores 975 new parts and 225 recovered parts. Total 1,200 parts (= 975+225) are shipped to module assembler. Module assembler assembles 1,200 parts and produces 1,200 new modules. Recovery center recovers the quality of the returned products and then sent 300 recovered modules to module inventory center. Module inventory center stores 1,200 new modules and 300 recovered modules. Total 1,500 modules are sent to product manufacturer for producing 1,500 products. 1500 products are sent to each retailer via distribution center. In RL, 1,500 products from all customers are returned to recovery center through collection center. 900 recovered products (= 60% × 1,500) of all returned products are resold at secondary markets. 75 unrecovered parts (= 5% × 1,500) of all returned products are sent to waste disposal center.

Based on the results of Table 4, Figs. 2 and 3, we can reach the following conclusions.

- The proposed CLSC model can represent the detailed material flows at each stage and effectively handle the recovered part and module by using two inventory centers, when compared with the conventional models of Amin and Zhang [1], Wang and Hsu [12], and Chen et al. [3].
- The a-HGA approach shows to be more efficient in many measures of performance than the GA, HGA and Lingo, which implies that the former can explore whole search space rather than the latter do.
- The search speed of the a-HGA is significantly slower than those of the GA and HGA, since the former has an adaptive scheme to regulate crossover and mutation operators and the search structure requires many times.

6 Conclusion

In this paper, we have proposed a new type of the CLSC model. The proposed CLC model has part suppliers at areas 1, 2, 3, and 4, module assembler, product manufacturer, distribution center and retailer for FL and customer, collection center, recovery center, secondary market and waste disposal center for RL. Especially, for effectively handling recovered part and module, two inventory centers (part and module inventory centers) have been used.

The proposed CLSC model has been represented by a mathematical formulation, which is to minimize the sum of handling cost, fixed cost and transportation cost resulting from each stage of FL and RL under satisfying various constraints. The mathematical formulation has been implemented by the a-HGA approach. The a-HGA approach is a hybrid algorithm with GA and CS approach. Also, using an adaptive scheme, the rates of crossover and mutation operators are automatically regulated in a-HGA approach. The a-HGA approach has been implemented in various scales of the proposed CLSC model to compare its performance with those of GA, HGA and Lingo. The experimental results have shown that the a-HGA approach is more efficient in terms of various measures of performance than the GA, HGA and Lingo. However, since the search speed of the a-HGA approach is significantly slower than those of the others, a room for improvement is still left in the a-HGA approach.

Acknowledgments. This work is partly supported by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS): No. 15K00357.

References

1. Amin SH, Zhang G (2012) An integrated model for closed-loop supply chain configuration and supplier selection: multi-objective approach. Expert Syst Appl 39(8):6782–6791
2. Amin SH, Zhang G (2013) A multi-objective facility location model for closed-loop supply chain network under uncertain demand and return. Appl Math Model 37(6):4165–4176
3. Chen YT, Chan FTS, Chung SH (2014) An integrated closed-loop supply chain model with location allocation problem and product recycling decisions. Int J Prod Res 53(10):3120–3140
4. Fleischmann M, Krikke HR et al (2000) A characterisation of logistics networks for product recovery. Omega 28(6):653–666
5. Gen M, Cheng R (1997) Genetic algorithms and engineering design. Wiley, New York
6. Gen M, Cheng R (2000) Genetic algorithms and engineering optimization. Wiley, New York
7. Georgiadis P, Besiou M (2008) Sustainability in electrical and electronic equipment closed-loop supply chains: a system dynamics approach. J Cleaner Prod 16(15):1665–1678
8. Kanagaraj G, Ponnambalam SG, Jawahar N (2013) A hybrid cuckoo search and genetic algorithm for reliability–redundancy allocation problems. Comput Ind Eng 66(4):1115–1124
9. Lingo (2015) Lindo Systems. www.lindo.com
10. Savaskan RC, Bhattacharya S, Van Wassenhove LN (2004) Closed-loop supply chain models with product remanufacturing. Manage Sci 50(2):239–252
11. Srinivas M, Patnaik LM (1994) Adaptive probabilities of crossover and mutation in genetic algorithms. IEEE Trans Syst Man Cybern 24(4):656–667
12. Wang HF, Hsu HW (2010) A closed-loop logistic model with a spanning-tree based genetic algorithm. Comput Oper Res 37(2):376–389

The Innovation Research of College Students' Academic Early-Warning Mechanism Under the Background of Big Data

Yu Li and Ye Zhang(✉)

University of Electronic Science and Technology,
Chengdu, People's Republic of China
zhangye_uestc@126.com

Abstract. With China's higher education from "elite education" to "mass education", students' academic problems have become increasingly serious. In the big data era, building college students' academic early-warning mechanism is of great significance for the training of qualified college students and improving the quality of higher education to keep pace with the times. However, in recently, the college students' research of academic early-warning mechanism based on the big data is relatively rare. This paper introduces the current research status of academic early-warning mechanism and the background of big data, and summarizes the characteristics of big data. Then it expounds the impact of big data thinking on the traditional academic early-warning mechanism about the four aspects: decision-making, forecasting, service and evaluation. This article focuses on putting forward an innovation model of academic early-warning mechanism concerning the information collection system, information processing and analysis system, program-making system and information release system based on big data, which will guide the practice of big data in the field of college students' academic early-warning.

Keywords: Academic early-warning · Big data · Impact · The innovation model

1 Introduction

After the reform and opening up, China's higher education has changed from elite education to mass education. The admission rate of college entrance examination has risen from 7% to 80% Since the 1980s, but the quality of students has been declining and the academic problems have been increasing year by year. For instance, the number of students who demote and dropped out of school increased obviously. Therefore, it is necessary to establish the academic early-warning mechanism to solve these problems.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_86

Academic early-warning is an impact on individual students in serious academic situation, and promptly reminds the students themselves and informs the parents. It is a academic early-warning system through the communication and collaboration of school, parents and college students, targeting to take appropriate preventive measures between the undergraduate students in the learning situation to give timely guidance to help students solve their academic problems and successfully complete their studies. In recent years, the research on college students' academic early-warning mechanism has lots of ideas. Chen [4] put forward the design idea of the colleges' early-warning mechanism from the aspects of the connotation, goal, function, warning system and the early warning level on account of the angle of the credit system of teaching management in colleges and universities. Cao [3] showed various reasons of academic problems behind the "learning difficulty students", putting forward the establishment of "learning difficulty students" assistance mechanism of academic pre-warning. Ding [5] and his colleagues took Science and Technology University in Suzhou province of China as an example, comparing and analyzing the data of school enrollment of academic early-warning mechanism. Hua [7] introduced the regulations and characteristics of academic early-warning system of Taiwan Universities in China, and then discuss the study of academic early-warning mechanism of college students in Mainland China. According to the document of Ministry of Education, which gives many opinions to improve the quality of teaching on further deepening reform of undergraduate teaching in 2007, it proposed to strengthen the management of personnel training process. "National Medium and Long Term Education Reform and Development Plan (2010–2020)" clearly published: establishing the assistance mechanism of students with learning difficulties, improving the excellent student training methods. In August 2007, the Ministry of Education promulgated "Academic Early-warning" as one of 171 new Chinese Words. The above studies and documents indicate that the system of academic early-warning is of great significance to the cultivation of talents in colleges and universities.

With the advent of the big data era, data increases fast, and the date range is expanded, which contribute to the formation of educational database. The information plan of national education offered "two platforms" (the public service platform of national education resources, the service platform of national education management) based on the concept of big data in wisdom campus construction and the large-scale opening online courses' teaching practice. In the whole sample data, mining the regulation of educational development is conducive to the guidance of education in colleges and universities to improve the accuracy of decision-making. To a certain degree, in the construction of educational information at different levels has taken a historic step, furthermore, it also lays the foundation for the practice of big date in the field of academic early-warning. However, although many scholars have explored the application of big data in education field at present, the theoretical and applied research on big data in the academic early-warning mechanism of college students in China and abroad is still in a preliminary stage. Therefore, this article will analyze the impact of big data on the traditional academic early-warning mechanism and design the model of college students' academic early-warning mechanism under the background of big data, and expect to provide basis for the follow-up research and application of big data in the field of college students' early-warning mechanism.

2 The Brief Introduction of Big Data

There is no uniform definition of big data at present, citing the McKinsey Global Institute's description of The Next Frontier in Innovation, Competition and Productivity: big data refers to the date process that is different from traditional software tools [8]. Although the definition of big data is not uniform, the characteristics of big data are clear and recognized. (1) The huge data volume. A variety of terminal equipment and sensors produce a lot of data, PB-scale data sets can be described as normal. (2) The variety data types. In the era of big data, more and more unstructured data, including network log, audio, video, pictures, geographical information. These different types of data needs higher requirements on data processing. (3) Processing speed. The most significant features is to distinguish big data technology from traditional data technology. In the face of massive amounts of complex data, big data can deal with real-time data faster. (4) Low value density. The value density is inversely proportional to the size of the data, for an hour video, the useful data may be only one or two seconds in the continuous monitoring.

The first feature of big data is large-scale, with the development of information technology, a variety of information systems, databases, cloud storage, Internet, Web of Things and mobile intelligent terminals, especially in recent years, increase rapidly in social system. The data sharing becomes very easy, so that the data scale is expanding constantly. At the same time, big data also has a complex and diverse data structures, owing to large, complex and volatile date, it is becoming more and more difficult to obtain hidden, useful knowledge, traditional data warehouse and data mining-related processing mode has been unable to meet the requirements, big data has a stronger decision-making power, insight and process optimization capabilities, which brings new changes to data processing, storage and conversion.

The latent value of big data is the correlation between the data, big data thinking is a conversion from the traditional causal analysis to the relevant analysis. More and more countries, governments, industries, enterprises and other institutions have realized that big data is becoming the organization's the most important asset, the capabilities of data analysis is becoming the organization's core competitiveness [11]. At present, the government has put big data into application to promote people's daily lives, the internet has penetrated into various industries. Big data has brought about an enormous effect on education. It has been integrated with education and is promoting the reform of educational model. Besides, it grounds for the application in college students' academic early-warning mechanism.

Big data is a prediction method based on objective data that has been obtained rather than on subjective inferences. This analysis method is based on the huge data, which is obviously different from the sample analysis method which has been used in the past. It can correct the error of the sampling method and improve the accuracy of the analysis result. Big data analysis can get the precise results, easier access to the user's trust, and thus has a wide utilization. The application of big data technology in college students' academic early-warning mechanism will have more timely and accurate early-warning.

3 The Impact of Big Data Thinking on the Traditional Academic Early-Warning Mechanism

3.1 Rational Decision-Making

The main object of universities and colleges is college students who as the most dynamic and most potential individuals, college students' thinking, behavior and personality are diverse. Due to the invisible and complex nature of thinking, it is hard to understand one's mind. In the past, we could only make judgments on the basis of individual experience and simple sample statistics. This traditional subjective decision-making method and the statistical analysis method of small sample data will play a certain role in the traditional academic pre-warning, but the cost of manpower, material and financial resources is huge, and the accuracy of academic early-warning needs to be improved. In the era of big data, we can effectively make more scientific judgments, more rational decision-making. Big data can provide a data basis for us to make the rational decision through the real-time monitoring of students learning and lives in and out of class. We can also analyze the students' thoughts and behavior through the internet, forming regular recognizing of them by the analysis of vast amounts of date to achieve scientific decision-making.

3.2 Precise Prediction

Forecast is the core of big data, the big data analysis method applied to the massive data, so as to predict the possibility of occurrence, to achieve the purpose of prediction. A large number of data not only allows us to predict the development of things, but also makes our prediction of human behavior possible. In the era of big data, the behavior data that is interdependent and associated of college students have been recorded and preserved. We can find the link between their behaviors and predict trends through the depth analysis and integration of behavioral data, so that we are able to react quickly to guide and intervene in advance.

3.3 Personalized Service

The big data has opened the way of individualized education, it can reveal the students' learning pattern through the data tracking and analysis of the students' learning process, formulating individualized educational programs for them, and giving the practical direction for individualized education. College students are highly individualized groups. They pay attention to strong character hoping to be treated as a unique individual. In the past, students worked only from the overall work programs, which ignore the personal differences and individual needs. Big data concerns every person, and will make a big change to our students work. It will make us re-examine the work of students not only from the overall picture, but also pay more attention to individual work to carry out specific work to promote individualized development of each college student. Through the whole, timely and dynamic record of each student's learning, life and social situation, the big data contribute to a better understanding personality and growth needs of students so that we can adopt ideological and political education measures, career planning, psychological counseling, comprehensive quality education to achieve personalized service of college students.

3.4 Scientific Evaluation

In the past, it is difficult to quantify the college students work whether it is the ideological evaluation of students or students' the economic situation of the family. It is inevitable that there is a certain bias resulting in a highly subjective evaluation that is because of the information only from counselors, class teachers, students and other channels. But the big data can promote the scientific evaluation of teaching to achieve more effective academic early-warning. To assess the economic situation of students in their family, for example, we can clearly grasp the specific period of time students' income and expenditure through their campus cards consumption records, shopping consumption records, mobile payment list, personal account records, thus we can make an accurate judgment of their personal economic situation as an important basis for judging the economic status of their families to avoid the errors caused by the subjective analysis. In the evaluation of the students' ideological situation, the analysis of the massive data can be more deterministic to test the dynamics of their thoughts and behavior to reflect the information of their ideological characteristics by data processing.

4 The Innovation Research: The Building of Academic Early-Warning Mechanism Under the Big Data Environment

In the era of big data, the academic early-warning mechanism consists of four major systems: information collection system, information analysis and processing system, program-making system and information release system.

4.1 Information Collection System

The collection of information is the primary work of the academic early-warning system, and it is the key to the work of academic early-warning as well. In the coming of big data, the basis of college students' work is to master the huge data which can really understand the characteristics of the behavior of college students to effectively find out the education, management and service countermeasure. First, colleges and universities should conduct the top-level design that is an overall vision of an integrated data platform inside their own campus. Considerable universities in the development of the wisdom campus's construction, different departments of the university only consider the work of the department needs, so that it is difficult to achieve data sharing and integration. So the school should set up a coordination department and academic pre-warning department to build a data center to gather the information of the office, academic affairs, school work, library and other related departments in order to form information platform. Secondly, building a system of online data collection platform to form a large database of students across the school in order to ensure the timely and complete collection of all students of all data. At the same time, colleges and universities should from the overall point of view do a good job of data classification, hierarchical collection and planning work to ensure the diversify of data sources and data types. Thirdly, colleges and universities should take the initiative to share the social database contributing to the establishment of big database of education. In this paper, the educational data can be divided into three categories: explicit data, behavioral data and system data (Fig. 1). Explicit data is the data input or output by the end user. Behavior data can also be called control data, which is designed by the developer for the purpose of recording the data of the operation process of the user, generally only seen by the administrator. System data is automatically generated by the system, behavioral data and system data are hidden data [9].

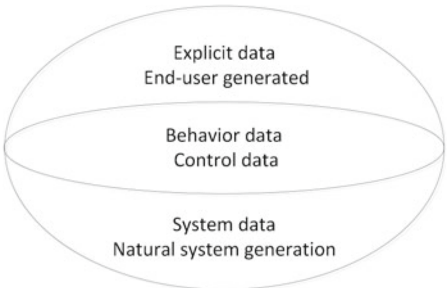


Fig. 1. The three categories educational data

4.2 Information Analysis and Processing System

The amount of information is large and have different structure, and a lot of information have no effect on the academic early-warning, so we need information processing and analysis system to sorting, deleting unnecessary information, and find useful details to making the academic early-warning mechanism more effective. The work of this system is to dynamically monitor the behavior of college students inside and outside the classroom, using the data mining technology of big data to analyze the collected information, and discover the abnormal phenomenon or behavior of college students. Under the big data environment, we use the technology of Outlier Mining Algorithm to find out the data away from the regular objects and launch the analysis of the outlier data. From the data itself, outlier data are deviations from normal data, which may be produced due to errors in the data formation process, and are often eliminated directly in early data analysis. However, due to data from reality, a lot of outlier data is not coming into being due to the errors, but the corresponding data source itself does contain a special behavior that may also mean a very important information that indicates the emergence of a new situation, which need us to distinguish, so the mining and analysis of outlier data is of great significance. The general idea of outlier mining is that: in a data set with n data objects, a desired outlier objects k ($k \leq n$) is given, and dig out a number of outliers data that are significantly different from the rest of the objects in the data set [1]. There are a number of outlier mining algorithms, including statistic-based, distance-based and bias-based detection algorithms, and outlier-based data using conventional data mining algorithms. Through scientific monitoring and analysis, we can timely identify the adverse factors, issue academic early-warning to avoid the depravation in academic.

4.3 Program-Making System

The Course Signals System divides students into three groups: green, yellow, and red according to the current state [2]. The green signal indicates that the students are likely to achieve the goal if he or she continues the current learning state. The yellow signal indicates that the student has potential difficulties in learning the courses, while the red signal indicates that students may fail in the exams. Under the big data background, we are able to output the visualization of the early-warning signals by using the technology of big data mining and analysis, so that students can continue to maintain the current status or make corresponding improvements.

(1) The early-warning program of the red signal: for red alert students, firstly, we need to analyze the reasons of their learning difficulties one by one, and arrange mutual help with each other. Counselor or class teacher would better chant with each students as much as possible to help them find causes of learning decline, and then seek out appropriate learning methods to improve academic grade. Secondly, for red warning students, we can be appropriate to reduce their extracurricular activities, and increase learning time, learning content. At the same time, we should not encourage them to take too much non-professional training curriculum.
(2) The early-warning program of the yellow signal: it is better to give encouragement to the yellow alert students. We could not only affirmed their existing achievements, but also put forward higher requirements. What's more, the

yellow warning for students can also be divided into two types to be treated in different. Some students study hard but not get good achievements. Their academic performance is not good, which is not caused by laziness. They take a lot of time to learn, but do not learn too much, we can arrange study-well students to give them one-on-one help, and offer guidance on learning methods. Other students lack initiative in learning, it is better to carry out strict management measures to require poor autonomy students. For example, we can take certain treatment measures, such as the centralized management, self-study at night, roll-call system.

(3) The early-warning program of the green signal: the green students are very excellent. We should pay attention to control their complacency, and put forward higher requirements to keep the existing achievements and continue to improve their comprehensive ability. Besides, we should ask them to learn well in the professional courses, doing well in the minor course at the same time, and inspire them to reach high standards in its computer level and English ability.

4.4 Information Release System of Academic Early-Warning

Release system is helpful for university students to understand their current academic situation by setting up information release system of academic early-warning. The school provides corresponding help for the students' academic difficulties. The parents strengthen supervision and guidance, it will achieve better warning effect through tripartite linkage among the school, family and students. On one hand, when the students log on to their academic early-warning system, the relevant early-warning information will automatically inform and highlight prompt, students can timely understand their own learning status by this confidential and practical warning, and they can form a certain psychological warning to take some self-remedial measures according to their specific circumstances. So we can set up release system by the method of abnormal affecting elements feedback [6].

Abnormal Affecting Elements Feedback
The key attribute subspace kas (p) of outlier object P is the main cause of data outliers, and can be used to accurately feedback the students' academic abnormalities with kas (p). Release system from two aspects of the q test results for output. Release system output test results of q times from tow aspects.

(1) Personalize abnormal elements analysis of outlier objects. Set the key attribute subspaces of the same outlier object p from every test are kas(p), $i = 1, 2, \cdots, p$. If P in the i time is not outlier, then kas(p) is an empty set. Set P is outlier in q_o, c_j is the occurring times of attribute a_j in all $kas(p)$. Apparently, the bigger is c_j, the more times of attribute a_j affecting P to outlier. Appraisal the affecting of attribute a_j to P abnormal study by c_j/p_o, to draw personalize abnormal elements affection chart. For example, one student is judged as abnormal 7 times in 10 tests data in one week, which

is more than 10/2, so he gets an orange warning and his personal affecting elements and their degree of affecting are: Job performance is 5/7, times of being late in class is 2/7, internet surfing duration is 4/7, downstream flow is 3/7, community activity duration is 3/7, staying in dormitory duration is 1/7, which shows the student's constitute of study abnormal elements by Fig. 2.

(2) Total degree of affecting analysis of element a_j, set the abnormal degrees of affecting of a_j in every test are c_j/k, $i = 1, 2, \cdots, q$. Draw elements continuously affecting chart to show the total situation of study abnormal affecting elements.

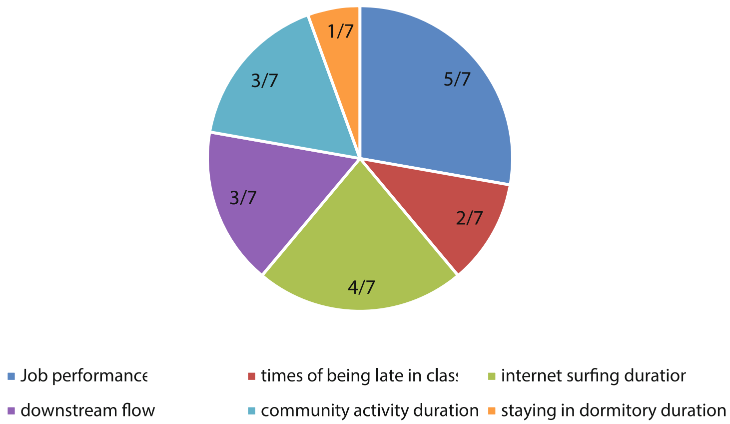


Fig. 2. The individual influence about abnormal factors

5 The Process Model of Academic Early-Warning Mechanism

In the big data environment, the main process of academic early-warning mechanism is information collection, information analysis, and information prediction and information release [10]. The process model of academic early-warning of college students is constructed based on the technology of big data. As shown in Fig. 3.

(1) The information collection system of academic early-warning: using the large database of information collection in university and the large database of education outside the university collects the internal and external information in colleges and universities, and using big data technology to initialize and classify them.

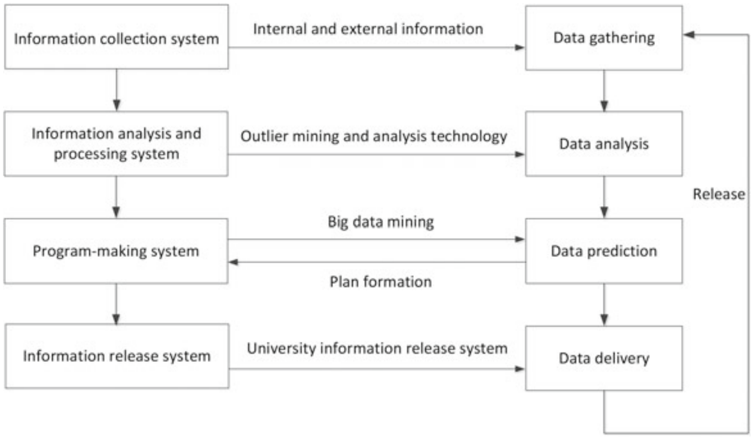


Fig. 3. The process model of academic early-warning of college students is constructed based on the technology of big data.

(2) The information analysis and processing system of academic early-warning: further cleaning and transforming to the collected information by using the big data analysis and processing technology, and using the outlier data mining and analysis technology of the big data to analyze the large number of data, focusing on the abnormal information and identifying adverse factors to make an early prevention.
(3) The program-making system of academic early-warning: according to Course Signals System, the early-warning of academic is divided into green, yellow and red pre-warning, and the big data is used to visualize the early-warning signals.
(4) The information release system of academic early-warning: college students get the feedback of results, and take measures to adjust learning state, which will provide feedback information to begin the next round of information collection.

6 Conclusion

Big data can collect data that is unable to obtain or acquire costly in the past, and more easily utilize their teaching level and management ability in the process of teaching analysis to fully enhance students' academic performance, and it will also set off a huge revolutionary by bringing influence in the field of education. However, the goal of education is the growth of each person that driving people to recognize things with a rational thinking, and gradually form a relatively perfect thinking of self-consciousness, which is so complex and hard to make perfect forecast. Therefore, on one hand, big data services for education in the academic early-warning system to be required cautious implementation in the practical application. On the other hand, it is so attractive and challenging to create personalized education through big data, more confidence should be given to design more detailed academic early-warning mechanisms to find more targeted interventions.

Acknowledgements. We would like to thank International Society of Management Science and Engineering Management and its workers. This article also benefit from the proposals offered by seminar participants of the 416 office at University of Electronic Science and Technology of China.

References

1. Annika W, Zdenek Z et al (2014) Improving retention: predicting at-risk students by analyzing clicking behavior in a virtual learning environment
2. Arnold KE, Pistilli MD (2012) Course signals at purdue: using learning analytics to increase student success. http://www.itap.purdue.edu/learning/does/research/ArnolejPistilli-Purdue-University-Course-Signals-2012.pdf2015-12-31
3. Cao W (2014) Analysis of the mechanism of the early-warning and aid for the students with learning disabilities. Higher Eng Educ Res 2014(5):93–96 (in Chinese)
4. Chen Q (2012) To explore the early warning mechanism of college students under the credit system. J Guangxi Teachers Educ Univ 2012(28):60–65 (in Chinese)
5. Ding F (2015) An empirical study on the performance of academic early-warning of college students. Frontiers Pract 2015(2):72–73 (in Chinese)
6. Ge D, Zhang S (2015) Educ Data Mining Methods Appl. Educational Science Press, Beijing
7. Hua J (2016) Learning early-warning system in Taiwan University. Educ Sci 2016(3):41–44 (in Chinese)
8. Manyika J, Chui M, et al (2011) Big data: the next frontier for innovation, competition, and productivity. Analytics pp 3–5
9. U.S. Department of Education Office of Educational Technology (2010) Transforming American education: learning powered by technology. http://www.ed.gov/sites/default/files/netp2010.pdf
10. Wei S (2016) Digging the value of education data in the era of big data. Mod Educ Technol 23(2):5–11 (in Chinese)
11. Witten IH, Frank E (2000) Data mining: practical machine learning tools and techniques with Java implementations. San Francisco: Morgan Kaufmann Publishers Inc

Bottleneck Management of Multi-stage Sorting-Packing Operations with Large-Scale Warehouse System

Tetsuya Sato(✉) and Hiroshi Katayama

Faculty of Science and Engineering, Waseda University, Tokyo 169-8555, Japan
tetsuya_sato@aoni.waseda.jp

Abstract. Integration of plural local DCs (Distribution Centers) into a large-scale DC with automatic warehouse system is considered as an effective strategy among logistics industries for responding to the diversifying needs of customers and markets, optimizing distribution operations and resource saving. However, as numerous items must be handled in such integrated DC, adequate design of the components such as warehouse, sorting-packing lines and their operation system is a critical issue for maintaining competitive advantage. This paper, focusing on the last issue, discusses the utility of TCM (Target Chasing Method) based dispatching policies for bottleneck management of entire process. Various PCVs (Pseudo Control Variables) are examined to utilized, which represent the state of the system, easy to observe but indirectly controllable (e.g. number of items and waiting time in queue, work load rate in each sorting-packing line). Performance analysis is examined by simulation experiments and obtained results suggest that smooth flow of items in each stage of sorting-packing line is realized by the proposed dispatching policies. Especially, dispatching policy for Just-In-Time delivery to customers that intends to adapt delivery time to customers, reveals the weak portion of the entire system.

Keywords: Automatic warehouse system · Sorting-packing operations · Multi-stage optimization · Heijunka production · Target chasing method · Just-in-time delivery

1 Introduction

In recent years, responding to the diversifying needs of customers and markets with optimizing distribution operations and resource saving are the trends among logistics industries. However, it requires to handle various types of products in small lots and the efficiency of each operation are decreased in conventional supply chain with plural local DCs (Distribution Centers). Thus, integration of local DCs into a large-scale DC with AWS (Automatic Warehouse System) is considered as an effective strategy for improvement of products flow in entire

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_87

supply chain. As numerous items must be handled in such integrated DC, adequate design of the components such as warehouse, sorting-packing lines and their operation system is a critical issue for maintaining competitive advantage.

For optimizing products flow in entire process of large-scale integrated DC, this paper discusses the utility of TCM (Target Chasing Method) [2,4–7] based dispatching policies. TOC (Theory of Constraints) [1] identifying and improving the bottleneck processes in multi-stage operations is an effective way for betterment of productivity. In this study, various PCVs (Pseudo Control Variables) [3,8] are examined to utilized, which represent the state of the system, easy to observe but indirectly controllable. Performance analysis is examined by simulation experiments and obtained results suggest that smooth flow of items in each stage of sorting-packing line is realized by the proposed dispatching policies. Especially, dispatching policy for Just-In-Time delivery to customers that intends to adapt delivery time to customers, reveals the weak portion of the entire system.

2 Experimental Method

2.1 Target Chasing Method (TCM)

TCM, which is frequently used in automotive industry, is a heuristical algorithm of sequencing method. Basically, TCM controls the order of job entry for minimizing the gap between ideal and real resource usage of the facility at each timing. Figure 1 shows the schematic diagram of TCM. Vertical and horizontal axes are the order of job entry and resource usage in the facility, respectively. From this figure, it is noticed TCM intends to pursue Heijunka (load levelling) production by controlling the order of job entry. The objective function of TCM is expressed in Eq. (1).

Objective Function:

$$\min\left(D_k\right) = \min_{i \in I} \sqrt{\sum_{j=1}^{R} \left(\frac{k \times N_j}{Q} - X_{jki} \right)^2}. \tag{1}$$

Symbols:

k : Order of job entry;
i : Lot number;
j : Resource type;
I : Set of the products;
R : Number of total resource;
Q : Total number of all products;
N_j : Usage of resource j for manufacturing all products;
X_{jki} : Usage of resource j for manufacturing from 1st to k^{th} product i;
D_k : The gap between ideal and real resource usage after k^{th} job entry.

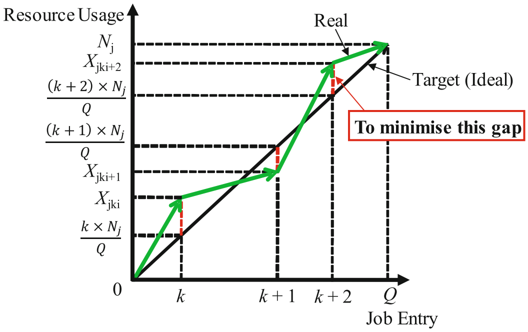


Fig. 1. Schematic diagram of target chasing method

2.2 Pseudo Control Variables (PCVs)

This paper focuses on the utility of PCVs [3,8], which represents the state of the system, easy to observe but indirectly controllable (e.g. number of items and waiting time in queue, work load rate in each sorting-packing line), for TCM-based dispatching policy. Figure 2 shows the schematic diagram of the controlled behavior of TCM using PCVs. Vertical and horizontal axes of Fig. 2 are the PCV values of each resource. In this case, there are two facilities in the same operational stage and the target of PCVs (diagonal line in Fig. 2) is defined as the same values between PCV1 and PCV2. Therefore, TCM controls current status of PCVs to the target line for Heijunka production. For example, the value of PCV2 is higher than PCV1 on the point P of Fig. 2, thus TCM tries to control PCV2 to decrease and PCV1 to increase for reducing the gap between actual and target PCVs.

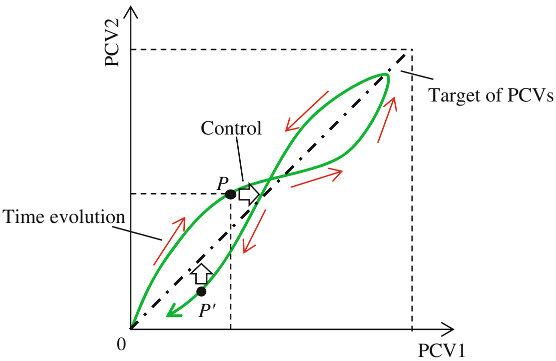


Fig. 2. Schematic diagram of target chasing method using pseudo control variables

2.3 Objective System

The objective system which is the multi-stage sorting-packing system in the large-scale DC with AWS is shown in Fig. 3. Dotted arrows in this figure are the routes of products flow from arrival to shipment. At first, handled products are arrived from plural local suppliers and warehoused to AWS. Products in the warehouse are sorted by AWS and dispatched to packing area. Products are re-installed to AWS after packing operation. This process is repeated in terms of multi-stage operation and packaged products are shipped to local assemblers. Table 1 shows the facilities and details of the operations in each stage of the objective DC. The 1st and 2nd stage of sorting and packing operations are the target to optimize by TCM based dispatching policies in this study, which is indicated by asterisks in Table 1. From Fig. 3, bottleneck management of the products flow at each packing station is required for efficient Just-In-Time delivery to the customers.

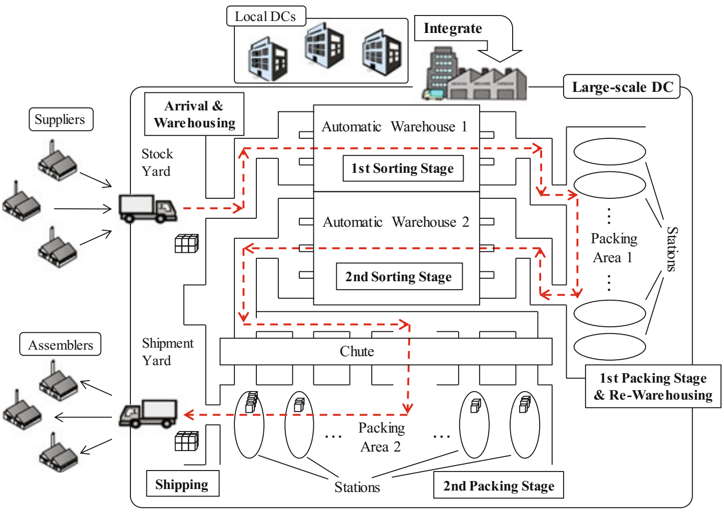


Fig. 3. Schematic diagram of objective large-scale distribution center

2.4 Simulation Model

This paper analyzes the performance of proposed TCM-based dispatching policies by simulation experiments for a downscaled problem, as original large-scale system is too complicated to tackle with. Figure 4 shows the schematic diagram of the defined simulation model. The buffers before each conveyor and AWS are inserted for dividing and identifying the bottlenecks of products flow in each stage.

Table 1. Details and facilities of the operations in each stage of the objective distribution center

Operation	Detail of operation	Facility
Arrival	Products arrive from suppliers	Stock Yard
Warehousing	Arrived products are stored to buckets and transferred to Automatic Warehouse	Stock Yard
1st sorting stage*	Warehouse system controls the dispatch order and the dispatched station of each product by the unit of a bucket	Automatic Warehouse 1
1st packing stage*	Products are extracted from the buckets and labeled. After labeling operation, the products are packed to inner boxes and re-stored to the buckets	Packing Area 1
Re-warehousing	Products are re-installed to Automatic Warehouse	Packing Area 1
2nd sorting stage*	Warehouse system controls the dispatch order and the dispatched station of each product by the unit of a bucket	Automatic Warehouse 2
2nd packing stage*	Required quantity of inner boxes with products are packaged to shipping boxes (or pallets)	Packing Area 2
Shipping	Products are shipped to assemblers	Shipment Yard

* Table 1 indicates the sub-processes of "Warehousing" and "Re-Warehousing"

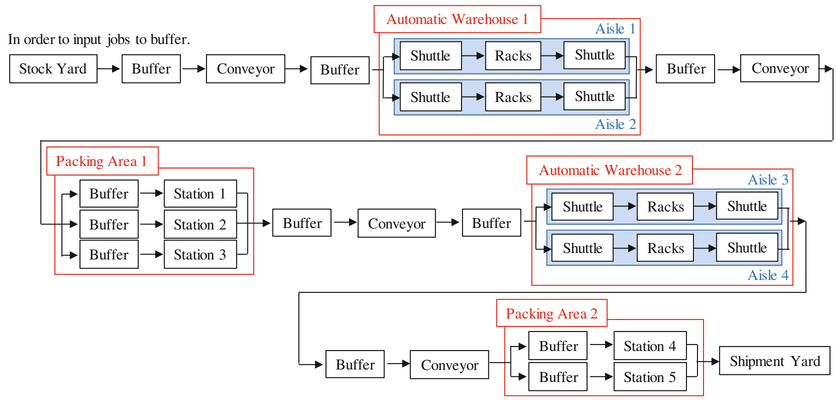


Fig. 4. Schematic diagram of the downscaled simulation model

The aim of this paper to analyze the performance of dispatching policies in AWS at each sorting stage; therefore, this paper unconsiders warehousing policy of AWS. The parameters of the simulation experiments are defined by properly deformed figures as follows:

- Number of product types: 3 (Product 1–3);
- Maximum load of the conveyors: 3 buckets;
- Supplying pitch of the conveyors: 1 bucket per unit time;
- Number of racks in each aisle: 100 racks;

Table 2. Facilities, processing time and lot size of each operation

Operation	Facility	Processing time	Lot size
1st sorting stage	Aisle 1 in Automatic Warehouse 1	$(2+3=)5^*$	1
	Aisle 2 in Automatic Warehouse 1	$(2+3=)5^*$	1
1st packing stage	Station 1 in Packing Area 1	3	1
	Station 2 in Packing Area 1	3	1
	Station 3 in Packing Area 1	4	1
2nd sorting stage	Aisle 3 in Automatic Warehouse 2	$(2+3=)5^*$	1
	Aisle 4 in Automatic Warehouse 2	$(2+3=)5^*$	1
2nd packing stage	Station 4 in Packing Area 2	15	6
	Station 5 in Packing Area 2	12	4

* Processing time of the aisles in each sorting stage are the sum of traveling times on the shuttles and the conveyors, respectively.

Table 3. Part of the generated lists of job entry and shipping order

(a) The list of job entry (Input)				(b) The list of shipping order (Output)				
Time of job entry	Product code	Station of packing Area 1	Station of packing Area 2	Shipping order	Product code	Lot size	Station of packing Area 1	Station of packing Area 2
1	1	3	4	1	3	4	2	5
2	2	3	4	2	1	4	1	5
3	1	3	4	3	1	6	2	4
4	2	3	4	4	2	4	2	5
5	1	3	4	5	2	4	2	5
6	3	3	5	6	3	6	2	4
7	2	3	4	7	2	6	1	4
8	3	3	5	8	2	4	1	5
9	3	3	5	9	2	6	3	4
10	3	3	4	10	3	6	2	4
11	3	2	4	11	3	6	1	4
12	2	2	5	12	3	6	3	4
13	3	2	5	13	2	6	3	4
14	2	2	5	14	1	4	1	5
15	1	2	4	15	3	6	2	4
...				...				
2000	1	2	5	410	2	6	2	4

- Buffer capacity of each packing station: 50 buckets;
- Number of job input to a queue: 2,000 products (expected as warehousing products in a week);
- Facilities, processing time and lot size of each operation: shown in Table 2.

The lists of job entry and shipping order of packaged products have been generated randomly in advance, which are shown in Table 3. Note that all products are assigned to operational stations in each packing stage beforehand of warehousing based on the job entry list which described the details of the packing operations in each station.

Performances of the proposed dispatching policies are analyzed by the number of items and waiting time in queue and work load rate in each sorting-packing line. Work load rate in sorting line (AWS) is defined as the rack occupancy. The definitions of waiting time and work load rate in packing lines are expressed in Eqs. (2) and (3), respectively.

Waiting time in queue:

$$QT_i^{x_j} = OT_i^{x_j} - IT_i^{x_j}. \tag{2}$$

Work load rate in packing line:

$$WR^{x_j} = \frac{\sum_{i=1}^{N^{x_j}} ST_i^{x_j}}{T}. \tag{3}$$

Symbols:

x_j : The j^{th} facility x;
OT_i^{xj} : Dispatched time of product i from facility x_j;
QT_i^{xj} : Waiting time of product i in facility x_j;
IT_i^{xj} : Job entry time of product i into facility x_j;
WR^xj : Work load rate in facility x_j;
T : Total processing time of entire operations;
N^{xj} : The number of products of which operation is finished in facility x_j.

The proposed TCM-based dispatching policies in each stage of sorting operation using AWS are shown in Table 4. The multi-stage operations are improved by a combination of the policies in the 1st and 2nd sorting stage. Regarding to the policy code $S - Xa$ or $S - Xb$, its components, i.e. S, X, a or b, represent the stage of the sorting operations, the index of the policies and number of items or waiting time in queue for PCVs, respectively. For example, the combination of policy codes 1-1a and 2-0 represents adoption of indicated dispatching rule to the 1st sorting stage and dispatching rule to the 2nd sorting stage described in Table 4, respectively.

3 Experimental Results and Discussion

3.1 Analysis of the Present Dispatching Policy (Benchmark)

At first, this paper analyzes the performance of the present dispatching policy that is used as a benchmark. Figure 5 shows the time series of the number of items in queue at each packing station, where the policies 1-0 and 2-0 are applied. As a result, the number of items in each queue are changed independently by the corresponding policies in each packing stage. Especially, from Fig. 5(a), it is observed that buffer capacities of each station that is defined as 50 buckets in Packing Area 1 tend to be fully occupied.

Table 4. Proposed TCM based dispatching policies for each stage sorting operations

1st Sorting stage		2nd Sorting stage		Notes
Policy	Detail	Policy	Detail	
1-0*:	<Benchmark> To dispatch the transactions by FCFS (First-Come, First-Served) principle	2-0*:	<Benchmark> To dispatch the transactions lot where all the items are ready to pack in the 2nd packing stage	In case of the order of shipping schedule is *not considered.*
1-1a:	To dispatch the transactions to the station by TCM, where number of items in queue is less than other stations at the 1st packing stage	2-1a:	To dispatch the transactions to the station by TCM, where number of items in queue is less than another station at the 2nd packing stage	
1-1b:	To dispatch the transactions to the station by TCM, where waiting time in queue is shorter than other stations at the 1st packing stage	2-1b:	To dispatch the transactions to the station by TCM, where waiting time in queue is shorter than another station at the 2nd packing stage	
1-2a:	To dispatch the transactions to the station by TCM with consideration of shipping schedule, where number of items in queue is less than other stations at the 1st packing stage	2-2a:	To dispatch the transactions to the station by TCM with consideration of shipping schedule, where number of items in queue is less than another station at the 2nd packing stage	In case of the order of shipping schedule is *considered*.
1-2b:	To dispatch the transactions to the station by TCM with consideration of shipping schedule, where waiting time in queue is shorter than other stations at the 1st packing stage	2-2b:	To dispatch the transactions to the station by TCM with consideration of shipping schedule, where waiting time in queue is shorter than another station at the 2nd packing stage	

* Current dispatching policies at each sorting stage.

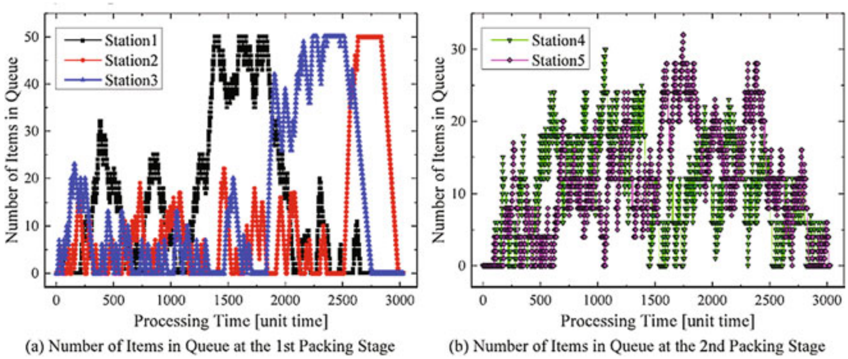


Fig. 5. The time series of the number of items in queue at each packing station in case of policies 1-0 and 2-0

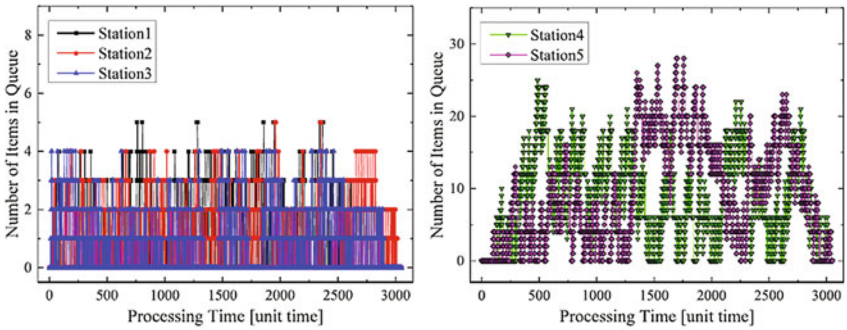
Table 5 shows the summary of criteria values in each operation in case of policies 1-0 and 2-0. The number of items and waiting time in queue of Packing Area 1 are both worse values than Packing Area 2 as well as the work load rate. Thus it is noticed that the 1st packing stage is the worst bottleneck operation in entire system.

Table 5. Averages and standard deviations of the criteria values in case of policies 1-0 and 2-0

	(a) Number of items in Queue [units]				(b) Waiting time in Queue [unit time]				(c) Work load rate			
	AWS1	Packing Area 1	AWS2	Packing Area2	AWS1	Packing Area 1	AWS2	Packing Area2	AWS1	Packing Area 1	AWS2	Packing Area2
Ave	0.661	38.681	44.011	21.503	1.000	57.524	30.803	31.534	0.33%	50.01%	22.01%	83.44%
SD	0.813	52.552	35.066	18.402	0.000	55.398	30.315	18.612	0.46%	49.94%	7.80%	37.17%

3.2 Bottleneck Management of the 1st Packing Stage

Based on the results in the last section, this paper analyzes the performance of proposed TCM-based dispatching policy applied to the worst bottleneck operation observed in the 1st packing stage. Figure 6 shows the time series of the number of items in queue at each packing station where the policies 1-1a and 2-0 are applied. In this case, policy 1-1a uses the number of items in queue as PCVs for TCM-based dispatching policy. From Fig. 6(a), the number of items in queue try to mutually equalize the same PCV values of other stations at the 1st packing stage by TCM-based dispatching. Note that the products are re-installed after the 1st packing operation and dispatched from AWS2 by policy 2-0, as in the last section. Thus, Fig. 6(b) are slightly changed from Fig. 5(b); nevertheless, the tendency is similar to the benchmark result.



(a) Number of Items in Queue at the 1st Packing Stage (b) Number of Items in Queue at the 2nd Packing Stage

Fig. 6. The time series of the number of items in queue at each packing station in case of policies 1-1a and 2-0

Table 6 shows the summary of criteria values in each operation in case of policies 1-1a and 2-0. The number of items and waiting time in queue at Packing Area 1 are extremely decreased from the benchmark result described in Sect. 3.1 and the work load rate in AWS1 is increased. For this reason, it is expected that the proposed TCM-based dispatching policies using PCVs facilitates Heijunka production effectively in the objective system.

In addition, policy code 1-1b uses the waiting time in queue as PCVs for TCM-based dispatching policy; however, the tendency of the criteria values is similar. Thus, this paper discusses only the cases that use the number of items in queue as PCVs in the following sections.

Table 6. Averages and standard deviations of the criteria values in case of policies 1-1a and 2-0

	(a) Number of items in Queue [units]				(b) Waiting time in Queue [unit time]				(c) Work load rate			
	AWS1	Packing Area 1	AWS2	Packing Area2	AWS1	Packing Area 1	AWS2	Packing Area2	AWS1	Packing Area 1	AWS2	Packing Area2
Ave	68.191	2.967	45.355	17.671	58.185	3.527	32.874	25.967	34.10%	49.58%	22.68%	82.73%
SD	72.045	3.912	36.089	15.591	75.736	3.209	34.045	17.020	21.77%	49.94%	7.73%	37.80%

3.3 Bottleneck Management of Entire System

From the benchmark result in Sect. 3.1, the second worst bottleneck operation in the system is identified, that is Packing Area 2. Due to this reason, this paper analyzes and discusses the proposed TCM-based policy for improving products flow of both the 1st and 2nd packing stages in this section. Figure 7 shows the time series of the number of items in queue at each packing station where the policies 1-1a and 2-1a are applied. From this figure, the number of items, which is defined as PCV in the policies, in each station at both the 1st and 2nd packing stages try to mutually equalize the same PCV values by TCM-based dispatching.

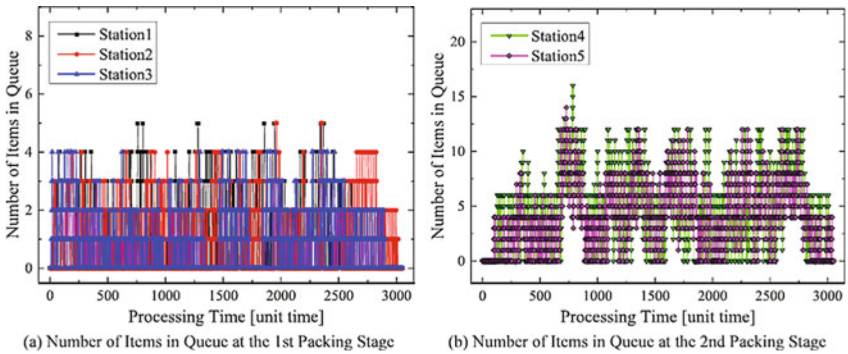


Fig. 7. The time series of the number of items in queue at each packing station in case of policies 1-1a and 2-1a

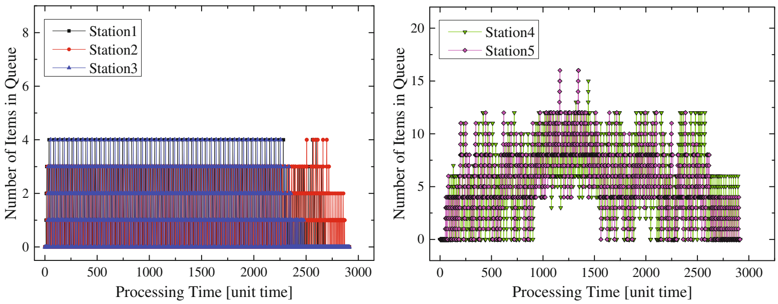
Table 7 shows the summary of criteria values in each operation in case of policies 1-1a and 2-1a. The number of items and waiting time in queue of the 2nd packing stage are decreased from the result of the last section, not only the 1st stage. Therefore, proposed TCM-based dispatching policies are utilized for bottleneck management with Heijunka concept of entire operation. On the other hand, the work load rate of Packing Area 2 in Tables 5, 6 and 7 are almost the same level. This is caused by the lot sizes and the processing time of each station at Packing Area 2 which are larger and longer than other processes, that is, after the improvement of this section, the bottleneck operation of the entire system, which is the 2nd packing stage, has not been transferred to other operations from the result of the last section. Thus, more improvement of operations throughput in Packing Area 2 is required for efficient improvement of the entire system, not only the preceding stages.

Table 7. Averages and standard deviations of the criteria values in case of policies 1-1a and 2-1a

	(a) Number of items in Queue [units]				(b) Waiting time in Queue [unit time]				(c) Work load rate			
	AWS1	Packing Area 1	AWS2	Packing Area2	AWS1	Packing Area 1	AWS2	Packing Area2	AWS1	Packing Area 1	AWS2	Packing Area2
Ave	68.191	2.967	54.340	8.358	58.285	3.527	39.689	11.755	34.10%	49.58%	27.17%	82.73%
SD	72.045	3.912	44.170	7.497	75.736	3.209	47.651	7.435	21.77%	49.94%	10.19%	37.80%

3.4 Bottleneck Management with Consideration of the Shipping Schedule

Finally, this paper analyzes bottleneck management with consideration of the shipping schedule. The assignment of all products to the operational stations at each packing stage in Table 3(a) decreases the degree of freedom in optimization of dispatching policy; therefore, product jamming is occurred by optimization with consideration of the shipping order. For this reason, the investigation of this section focuses the utility of dynamic reassignment of the products to each operational station. Figure 8 shows the time series of the number of items in queue at each packing station considering scheduled shipping order as Table 3(b), where the policies 1-2a and 2-2a are applied. As a result, the number of items in each station at both the 1st and 2nd packing stages try to mutually equalize the same PCV values, which is similar to the analyzed result of the last section.



(a) Number of Items in Queue at the 1st Packing Stage (b) Number of Items in Queue at the 2nd Packing Stage

Fig. 8. The time series of the number of items in queue at each packing station in case of policies 1-2a and 2-2a

Table 8 shows the criteria values in each stage in case of policies 1-2a and 2-2a. In addition, the proposed dispatching policy dependent characteristics of the number of items in queue, waiting time in queue and the work load rate at each stage are summarized in Fig. 9. The bar and line charts in this figure are the averages and standard deviations of the criteria values, respectively. From Table 8, the work load rate in AWS 1 is decreased from the result of the last section shown in Table 7(c); nevertheless, the number of the items and the waiting time in queue of AWS 1 are extremely decreased from Table 7(a) and (b). This is

Table 8. Averages and standard deviations of criteria values in case of policies 1-2a and 2-2a

	(a) Number of items in Queue [units]				(b) Waiting time in Queue [unit time]				(c) Work load rate			
	AWS1	Packing Area 1	AWS2	Packing Area2	AWS1	Packing Area 1	AWS2	Packing Area2	AWS1	Packing Area 1	AWS2	Packing Area2
Ave	4.807	3.554	50.920	10.444	5.154	4.180	36.578	14.222	2.40%	51.92%	25.46%	86.62%
SD	11.048	4.554	46.495	8.753	20.815	3.276	74.151	7.768	7.18%	49.90%	12.27%	34.04%

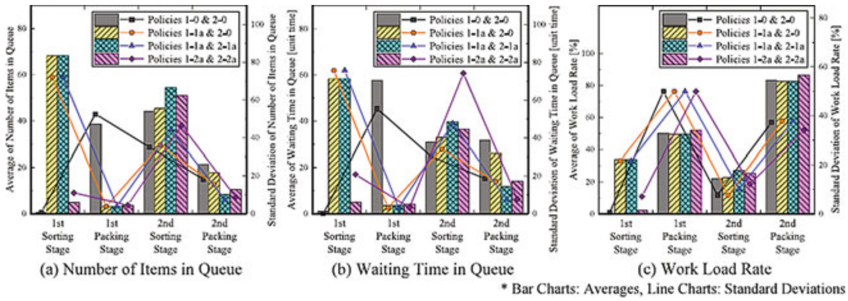


Fig. 9. Dispatching policies dependent characteristics of criteria values at each stage

caused by efficient dispatching is realized by dynamic reassignment of products to each operational station at the 1st packing stage. In contrast, the criteria value of AWS 2 and Packing Area 2 in Table 8 are almost the same level as in Table 7. This is caused by the lot sizes and the processing time of each station at Packing Area 2 which are larger and longer than other processes as similar to the last section. Therefore, proposed TCM-based dispatching policies with dynamic reassignment are utilized for efficient bottleneck management with Heijunka production of the entire operation.

4 Concluding Remarks

This paper examined the performance and the effectiveness of the proposed TCM-based dispatching policy using PCVs for bottleneck management of multi-stage sorting-packing operations in large-scale AWS-based DC. From the results, Heijunka production in each packing stage is expected to realize by the proposed dispatching policies. Moreover, this paper analyzed the case of given scheduled shipping order. In this case, dynamic reassignment of the transactions to each operational station at the 1st and 2nd packing stage facilitates optimization of dispatching operation effectively. Therefore, the proposed TCM-based dispatching policies with dynamic reassignment are utilized for effective bottleneck management for realizing Heijunka operation of the entire system.

The aim of this paper is the performance analysis of dispatching policies at each sorting stage in AWS for its relevant installation. Heijunka of the work load rate of each aisle at multi-stage AWS and operations at each stage are required for further improvement of throughput; which is a possible subject for future analysis.

Acknowledgements. This work was partially supported by JSPS KAKENHI (Grants-in-Aid for Scientific Research of Japan Society for the Promotion of Science) Grant Number JP16K17199.

References

1. Goldratt EM, Cox J (1984) The goal: a process of ongoing improvement. The North River Press, Great Barrington
2. Hwang R, Katayama H (2009) Integrated procedure of balancing and sequencing for mixed-model assembly lines: a multi-objective evolutionary approach. Int J Prod Res 40(21):6417–6441
3. Ishikawa S, Murata K, et al. (2012) Bottleneck management on multi-process sorting-packing system in distribution centre TCM implementation and its applicability reinforcement. In: Proceedings of The 15th J.S.L.S. national conference. The Japan society of logistics systems, 2nd–3rd June, pp 143–148 (in Japanese)
4. Katayama H (2006) Lean technology and its extensions-a case of target chasing method. In: Proceedings of the 11th Cambridge symposium on international manufacturing (CSIM 2006), Institute for Manufacturing, University of Cambridge, UK, 27th–28th September, 15 pages in CD-ROM
5. Katayama H, Tanaka M (1996) Target chasing method for work load stabilisation of production line-some advances of lean management technology. In: Proceedings of the 20th international conference on computers & industrial engineering (20th ICC & IE). Korean Institute of Industrial Engineers, Kyongju, 6th–9th October, pp 425–428
6. Kotani S (1983) Sequencing algorithms for the mixed-model assembly line. Toyota Technol 33(1):31–38 (in Japanese)
7. Monden Y (1983) Toyota production system: an integrated approach to just-in-time, 2nd edn. Industrial Engineering & Management Press, Norcross
8. Sato T, Ishikawa S, Katayama H (2016) Bottleneck management on multi-process sorting-packing system in large-scale distribution centre. In: Proceedings of the 2016 annual autumn meeting of japan industrial management association. Japan Industrial Management Association, Tokyo, pp 188–189 (in Japanese)

A Carbon-Constrained Supply Chain Planning Model

Zhimiao Tao and Jing Xu(✉)

Uncertainty Decision-Making Laboratory, Sichuan University,
Chengdu 610064, People's Republic of China
taozhimiao@scu.edu.cn, 20210431@qq.com

Abstract. This paper focuses on a simplified supply chain planning network consisting of manufacture and distribution in the context of carbon emission constraint. The problem is formulated as a bilevel programming model mathematically. An effective solution strategy is then presented to handle the bilevel model. A numerical example is proposed to illustrate the proposed model and algorithms. This study highlights the impact of carbon emissions constraint on the operational decisions of supply chain.

Keywords: Bilevel programming model · Supply chain planning · Carbon emissions · Carbon footprint

1 Introduction

More evidences show that human activity leads to global climate change via carbon emissions. Curbing the carbon emissions and controlling climate change have become a great challenge that all humanity has to face. As the main carbon emitters, firms must take their carbon emissions seriously.

A conventional thinking to curb the carbon emissions is to improve energy efficiency and install emissions control devices. However, such initiatives require capital investments. Meanwhile, efficient business practices and operational policies are another way to change carbon emissions. Different manufacture or order frequency may result in different emissions. Obviously, this way is more cost-effective, in contrast to the way of investing novel equipments. Some researchers have focused on the strategy to reduce the carbon emission through supply chains. Benjaafar et al. [5] integrated the carbon emission concerns into the operational decision-making with regard to procurement, production, and inventory management. They provided a series of operational models under different regulatory policies. Lee [15] used empirical case study of HMC (Hyundai Motor Company) and its key supplier's front bumper product to illustrate the issue of carbon footprint in supply chain management. Absi et al. [1] modeled the integration of carbon emission constraints in lot-sizing problems. Helmrich et al. [11] showed that lot-sizing with an emission capacity constraint is NP-hard and proposed several solution methods. Hua et al. [13] investigated how firms manage carbon footprints in inventory management under the carbon emission trading mechanism. Chen et al. [9] provided analytical support for the notion that it may be possible, via

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_88

operational adjustments alone, to significantly reduce emissions without significantly increasing cost. He et al. [10] examined the production lot0sizing issues of a firm under cap-and-trade and carbon tax regulations, respectively.

All the models in the literature mentioned above assume that there is a only decision maker and formulate as the problems as single level models. Single level decision-making structures don't reflect the real supply chain management. A supply chain naturally involves multiple decision components, which are collected in a hierarchical way. Individual activities are often governed by separate supply chain components which have their own, often mutually conflicting, objectives [7,17,19]. In addition, different participant components in the supply chain network may not operate with the same level of information. Some may possess more information, while other may possess less information, this may lead to information distortion [12,22,23].

Using a bilevel programming model, we study how carbon emission concerns can be integrated into supply chain planning. We examine how the values of emission control policies affect cost and emissions. An interactive approach is proposed to solve the model.

The rest of this paper is organized as follows. In Sect. 2, we make a problem statement as the motivation. In Sect. 3, we formulate a bi-level programming model for the supply chain planning problem under with strict carbon caps. An interactive approach is depicted in Sect. 4. In Sect. 5, we use the model to obtain managerial insights from the numerical experiment and discuss the implications of these insights to management practice and to public policy making. We summarize key findings and offer ideas for future research in Sect. 6.

2 Problem Statement

A simple supply chain normally consists of several manufacturing plants and one warehouse or distribution centre (DC), with several products (see Fig. 1).

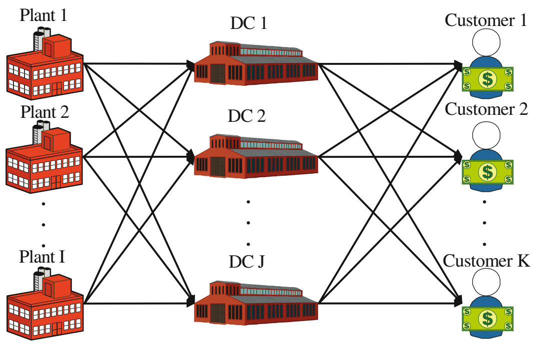


Fig. 1. A supply chain network

Given there is common decision maker who manages the whole supply chain, the supply chain planing problem can be formulated as an single level decision-making model. However, An alternative, more realistic, way to pose this problem is by recognizing: (i) the natural hierarchy that exists between the production and the distribution part of the supply chain and (ii) the fact that full information against the distribution part may not be really available at the production part and vice versa.

Without carbon emission considerations, the plants and the DCs determine the production amount and the inventory amount, respectively. Their objective In the presence of carbon emission considerations, the plants and the DCs must account for the emissions associated with various decisions regarding ordering, production, and inventory holding. In order to develop the mathematical model, we make some condition assumptions as follows:

(1) The emissions are linearly increasing in the associated decision variables;
(2) All the plant belongs one owner and all the DCs belongs to the other owner. DCs' owner has priority in decision;
(3) Customers' demands are stochastic;
(4) There is only one period in a planning horizon.

3 Model Formulation

In this section, we develop a mathematical model for the problem described above.

3.1 Notations

Before formulating the mathematical model, the needed notations are listed as follows.

Indices:
i: Plant $(1,\cdots,I)$;
j: DC $(j,\cdots,J)$;
k: Customer $(1,\cdots,K)$;
l: Product $(1,\cdots,L)$.

Parameters:
α_{il}: Capacity coefficient of product l at plant i;
β_{il}: Resource coefficient of product l at plant i;
γ_{jl}: Resource coefficient of product l at plant j;
a_{il}: Production cost coefficient for product l at plant i;
b_{ijl}: Transportation cost coefficient for product l from plant i to DC j;
h_{jkl}: Inventory holding cost coefficient for product l at DC j for customer k;
tr_{jkl}: Transportation cost coefficient for product l from DC j to customer k;
$\tilde{a}_{il}$: Emission coefficient due to production for product l at plant i;
$\tilde{b}_{ijl}$: Emission coefficient due to transportation for product l from plant i to DC j;
$\tilde{h}_{jkl}$: Emission coefficient due to inventory for product l at DC j for customer k;
$\widetilde{tr}_{jkl}$: Emission coefficient due to transportation for product l from DC j to customer k;
$\widehat{M}_{kl}$: Demand of product l at customer k;
P_i: Production capacity of plant i;
Q: Resources available to all the plants;

C_P: Emission cap for plants;
C_D: Emission cap for DCs.

Decision variable:
Y_{ijl}: Production amount of l at plant j for DC j;
X_{jkl}: Inventory of product l at DC j for customer k.

3.2 Modelling

For the owner of the plants, his (or her) to minimize the costs, which consists of its manufacturing cost and distribution cost between plants and DCs. Hence the objective of the production part Z_{PC} can be developed as

$$\min_{Y_{ijl}} Z_{PC} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L} a_{il}Y_{ijl} + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L} b_{ijl}Y_{ijl}. \tag{1}$$

We consider the strict emission cap regulation, the cap on carbon emissions of production part cannot be exceeded, i.e.,

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L} \tilde{a}_{il}Y_{ijl} + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L} \tilde{b}_{ijl}Y_{ijl} \le C_P. \tag{2}$$

Since production amounts from the plants should meet the levels required at the DC, we have

$$\sum_{i=1}^{I} Y_{ijl} \ge \sum_{k=1}^{K} X_{jki}, \quad \forall j, l. \tag{3}$$

The plant capacities are formulated as

$$\sum_{l=1}^{L}\sum_{j=1}^{J} a_{il}Y_{ijl} \le P_i, \quad \forall i. \tag{4}$$

Common used resources may be shared by all the plants, that is

$$\sum_{l=1}^{L}\sum_{i=1}^{I}\sum_{j=1}^{J} \beta_{il}Y_{ijl} \le Q. \tag{5}$$

For the distribution part, the decision maker's objective is to minimize the costs, then

$$\min_{X_{jkl}} Z_{DC} = \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L} h_{jkl}X_{jkl} + \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L} tr_{jkl}X_{jkl}, \tag{6}$$

in which the first term denotes inventory holding cost including material handling cost at DCs and the second indicates transportation cost from warehouses to customers. The carbon emissions from the distribution part also satisfies the emission cap, so

$$\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L} \tilde{h}_{jkl}X_{jkl} + \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L} \widetilde{tr}_{jkl}X_{jkl} \le C_D. \tag{7}$$

Sums of individual DCs' holding should meet demands of customers, that is

$$\sum_{j=1}^{J} X_{jkl} \geq \widehat{M}_{kl}, \ \ \forall k,l. \tag{8}$$

Demands' constraints (8) are stochastic types. In other words, we cannot ensure that the constraints are satisfied before the exact values of $\widehat{M}_{kl}$ are known. Some techniques have been developed to handle the stochastic case like this. Here, the chance-constrained technique proposed by Charnes and Cooper [8] is applied. Let θ_{kl} denote prescribed confidence levels. Applying chance-constrained technique to (8), the corresponding constraints are

$$Pr\left\{\sum_{j=1}^{J} X_{jkl} \geq \widehat{M}_{kl}\right\} \geq \theta_{kl}, \ \ \forall k,l, \tag{9}$$

where Pr means *probability*.

Inventory levels are limited by individual DC capacities, so

$$\sum_{k=1}^{K}\sum_{l=1}^{L} \gamma_{kl} X_{jkl} \leq R_j, \ \ \forall j. \tag{10}$$

In many cases, the decision maker has priority decisions on the production part are affected by parameters which are decided by the distribution part. For instance, production levels are decided from given information regarding the inventory conditions. Thus the supply chain planning model can be posed as the following bilevel decision-making problem:

$$\left\{\begin{array}{l}
\min\limits_{X_{jkl}} Z_{DC} = \sum\limits_{j=1}^{J}\sum\limits_{k=1}^{K}\sum\limits_{l=1}^{L} h_{jkl}X_{jkl} + \sum\limits_{j=1}^{J}\sum\limits_{k=1}^{K}\sum\limits_{l=1}^{L} tr_{jkl}X_{jkl} \\
\text{s.t.} \left\{\begin{array}{l}
\sum\limits_{j=1}^{J}\sum\limits_{k=1}^{K}\sum\limits_{l=1}^{L} \tilde{h}_{jkl}X_{jkl} + \sum\limits_{j=1}^{J}\sum\limits_{k=1}^{K}\sum\limits_{l=1}^{L} \widetilde{tr}_{jkl}X_{jkl} \leq C_D \\
Pr\left\{\sum\limits_{j=1}^{J} X_{jkl} \geq \widehat{M}_{kl}\right\} \geq \theta_{kl}, \ \ \forall k,l \\
\sum\limits_{k=1}^{K}\sum\limits_{l=1}^{L} \gamma_{kl}X_{jkl} \leq R_j
\end{array}\right. \\
\text{where } Y_{ijl} \text{ solves} \\
\left\{\begin{array}{l}
\min\limits_{Y_{ijl}} Z_{PC} = \sum\limits_{i=1}^{I}\sum\limits_{j=1}^{J}\sum\limits_{l=1}^{L} a_{il}Y_{ijl} + \sum\limits_{i=1}^{I}\sum\limits_{j=1}^{J}\sum\limits_{l=1}^{L} b_{ijl}Y_{ijl} \\
\text{s.t.} \left\{\begin{array}{l}
\sum\limits_{i=1}^{I}\sum\limits_{j=1}^{J}\sum\limits_{l=1}^{L} \tilde{a}_{il}Y_{ijl} + \sum\limits_{i=1}^{I}\sum\limits_{j=1}^{J}\sum\limits_{l=1}^{L} \tilde{b}_{ijl}Y_{ijl} \leq C_P \\
\sum\limits_{i=1}^{I} Y_{ijl} \geq \sum\limits_{k=1}^{K} X_{jki}, \ \ \forall j,l \\
\sum\limits_{l=1}^{L}\sum\limits_{j=1}^{J} a_{il}Y_{ijl} \leq P_i, \ \ \forall i \\
\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{I}\sum\limits_{j=1}^{J} \beta_{il}Y_{ijl} \leq Q \\
X_{jkl} \geq 0, \ \ \forall j,k,l \\
Y_{ijl} \geq 0, \ \ \forall i,j,l.
\end{array}\right.
\end{array}\right.
\end{array}\right. \tag{11}$$

4 Algorithm

Algorithms are the bridges between the theoretical models and the practical applications. The NP-hardness of bilevel programming problems make it full of challenges to design efficient algorithms. Researchers have proposed many approaches to solve bilevel programming problems [2, 4, 6, 14, 16, 18]. Among these algorithms, the interactive fuzzy programming methods are popular in recent years [3, 20, 24].

Model (11) can be reformulated in a general bilevel programming framework as follows:

$$\begin{cases} \max\limits_{\boldsymbol{x}\in R^{n_1}} F(\boldsymbol{x},\boldsymbol{y}) \\ \text{s.t.} \begin{cases} G(\boldsymbol{x},\boldsymbol{y}) \leq 0 \\ \text{where } \boldsymbol{y} \text{ solves:} \\ \begin{cases} \max\limits_{\boldsymbol{y}\in R^{n_2}} f(\boldsymbol{x},\boldsymbol{y}) \\ \text{s.t. } \ g(\boldsymbol{x},\boldsymbol{y}) \leq 0, \end{cases} \end{cases} \end{cases} \tag{12}$$

where, $\boldsymbol{x} \in R^{n_1}$ is the decision vector for the upper-level decision maker (leader) and $\boldsymbol{y} \in R^{n_2}$ is decision variable for the lower-level decision maker (follower); $F(\boldsymbol{x},\boldsymbol{y})$ is the objective function of the upper-level model and $f(\boldsymbol{x},\boldsymbol{y})$ is objective function of the lower-level model; $G(\boldsymbol{x},\boldsymbol{y})$ is the constraint of the upper-level programming and $g(\boldsymbol{x},\boldsymbol{y})$ is the constraint of the lower-level model.

The key idea is that the decision makers at both levels consider individual objectives separately as the initialization. Let $D = \{\boldsymbol{x},\boldsymbol{y} | G(\boldsymbol{x},\boldsymbol{y}) \leq 0, g(\boldsymbol{x},\boldsymbol{y}) \leq 0\}$. For each of the objective functions, assume that the decision makers have fuzzy goals such as "the objective should be substantially less than or equal to some value". The individual best values

$$F^{\min} = \min_{(\boldsymbol{x},\boldsymbol{y})\in D} F(\boldsymbol{x},\boldsymbol{y}),\ f^{\min} = \min_{(\boldsymbol{x},\boldsymbol{y})\in D} f(\boldsymbol{x},\boldsymbol{y})$$

and the individual worst values

$$F^{\max} = \max_{(\boldsymbol{x},\boldsymbol{y})\in D} F(\boldsymbol{x},\boldsymbol{y}),\ f^{\max} = \max_{(\boldsymbol{x},\boldsymbol{y})\in D} f(\boldsymbol{x},\boldsymbol{y})$$

are referred to when the decision makers elicit membership functions prescribing the fuzzy goals for the objective functions. The decision makers determine the membership functions $\mu_1(F(\boldsymbol{x},\boldsymbol{y}))$ and $\mu_2(f(\boldsymbol{x},\boldsymbol{y}))$ which are strictly monotone decreasing for objective functions, consulting the variation ratio of degree of satisfaction in the interval between the individual best values and the individual values. The domains of the membership functions of leader and follower are the interval $[F^{\min}, F^{\min}]$ and $[f^{\min}, f^{\min}]$, respectively. For the sake of simplicity, in this paper, we adopt a linear membership function, which characterizes the fuzzy goal of the decision maker at each level. The corresponding linear membership function $\mu_1(F(\boldsymbol{x},\boldsymbol{y}))$ and $\mu_2(f(\boldsymbol{x},\boldsymbol{y}))$ are defined as:

$$\mu_1(F(\boldsymbol{x},\boldsymbol{y})) = \begin{cases} 0 & \text{if } F(X_1,X_2) > F^{\max} \\ \frac{F^{\max}-F(\boldsymbol{x},\boldsymbol{y})}{F^{\max}-F^{\min}} & \text{if } F^{\min} \leq F(\boldsymbol{x},\boldsymbol{y}) \leq F^{\max} \\ 1 & \text{if } F(\boldsymbol{x},\boldsymbol{y}) < F^{\min}, \end{cases} \tag{13}$$

and

$$\mu_2(f(\boldsymbol{x},\boldsymbol{y})) = \begin{cases} 0 & \text{if } f(X_1,X_2) > f^{\max} \\ \frac{f^{\max}-f(\boldsymbol{x},\boldsymbol{y})}{f^{\max}-f^{\min}} & \text{if } f^{\min} \le f(\boldsymbol{x},\boldsymbol{y}) \le f^{\max} \\ 1 & \text{if } f(\boldsymbol{x},\boldsymbol{y}) < f^{\min}. \end{cases} \tag{14}$$

After eliciting the membership functions, leader subjectively specifies a minimal satisfactory level $\hat{\delta} \in [0,1]$ for his/her membership function $\mu_1(F(\boldsymbol{x},\boldsymbol{y}))$. Then, follower maximize his/her membership function subject to the condition that leader's membership function $\mu_1(F(\boldsymbol{x},\boldsymbol{y}))$ is larger than or equal to $\hat{\delta}$ under the given constraints, that is, followers solves the following problem:

$$\begin{cases} \max\limits_{\boldsymbol{x} \in R^{n_1}} \mu_2(f(\boldsymbol{x},\boldsymbol{y})) \\ \text{s.t.} \begin{cases} \mu_1(F(\boldsymbol{x},\boldsymbol{y})) \ge \hat{\delta} \\ (\boldsymbol{x},\boldsymbol{y}) \in D. \end{cases} \end{cases} \tag{15}$$

If an optimal solution to problem (15) exists, it follows that DM1 obtains a satisfactory solution having a satisfactory degree larger than or equal to the minimal satisfactory level specified by leader's own self. However, the larger the minimal satisfactory level is assessed, the smaller follower's satisfactory degree becomes. Consequently, a relative difference between the satisfactory degrees of leader and follower becomes larger and it is feared that overall satisfactory balance between both levels cannot maintain.

To take account of overall satisfactory balance between both levels, leader needs to compromise with follower on leader's minimal satisfactory level. To do so, a satisfactory degree of both decision makers is defined as

$$\lambda = \min(\mu_1(F(\boldsymbol{x},\boldsymbol{y})), \mu_2(f(\boldsymbol{x},\boldsymbol{y}))).$$

and the following problem is substituted for problem (15):

$$\begin{cases} \max \lambda \\ \text{s.t.} \begin{cases} \mu_1(F(\boldsymbol{x},\boldsymbol{y})) \ge \hat{\delta} \ge \lambda \\ \mu_2(f(\boldsymbol{x},\boldsymbol{y})) \ge \lambda \\ \lambda \in [0,1] \\ (\boldsymbol{x},\boldsymbol{y}) \in D. \end{cases} \end{cases} \tag{16}$$

For problem (16), consider the auxiliary problem

$$\begin{cases} \max \lambda \\ \text{s.t.} \begin{cases} \mu_1(F(\boldsymbol{x},\boldsymbol{y})) \ge \lambda \\ \mu_2(f(\boldsymbol{x},\boldsymbol{y})) \ge \lambda \\ \lambda \in [0,1] \\ (\boldsymbol{x},\boldsymbol{y}) \in D. \end{cases} \end{cases} \tag{17}$$

By solving problem (17), we obtain a solution maximizing a smaller satisfactory degree between those of both decision makers.

If an optimal solution $(\boldsymbol{x}^*,\boldsymbol{y}^*)$ to problem (17) satisfies the condition that $\mu_1(F(\boldsymbol{x}^*,\boldsymbol{y}^*)) \ge \hat{\delta}$, it follows that leader obtains a satisfactory solution. However, the solution $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ does not always satisfy the condition. Then the ratio of satisfactory degree between both levels

$$\Delta = \frac{\mu_2(f(\boldsymbol{x}^*, \boldsymbol{y}^*))}{\mu_1(F(\boldsymbol{x}^*, \boldsymbol{y}^*))} \tag{18}$$

is adopted.

If $\Delta > 1$, i.e., $\mu_2(f(\boldsymbol{x}^*, \boldsymbol{y}^*) > \mu_1(F(\boldsymbol{x}^*, \boldsymbol{y}^*))$, then leader updates the minimal satisfactory level $\hat{\delta}$ by increasing its value. Receiving the updated level $\hat{\delta}'$, follower solves problem (15) with $\hat{\delta}'$ and then leader obtains a larger satisfactory degree and follower accepts a smaller satisfactory degree. Conversely, if $\Delta < 1$, i.e., $\mu_2(f(\boldsymbol{x}^*, \boldsymbol{y}^*) < \mu_1(F(\boldsymbol{x}^*, \boldsymbol{y}^*))$, then leader updates the minimal satisfactory level $\hat{\delta}'$ by decreasing its value, and leader obtains a smaller satisfactory degree and follower accepts a larger satisfactory degree. The interactive process terminates when two conditions hold [21].

5 Numerical Example

Consider supply chain consisting of three plants, two DCs, three customers, with two products. The values of parameters are listed in Table 1. The demands follow the normal distribution, defined as

Table 1. Parameter values

$\alpha_{11} = 4$	$\alpha_{12} = 3$	$\alpha_{21} = 5$	$\alpha_{22} = 2$	$\alpha_{31} = 3$	$\alpha_{32} = 3$
$\beta_{11} = 7$	$\beta_{12} = 10$	$\beta_{21} = 12$	$\beta_{22} = 6$	$\beta_{31} = 8$	$\beta_{32} = 9$
$\gamma_{11} = 6$	$\gamma_{12} = 8$	$\gamma_{21} = 10$	$\gamma_{22} = 4$		
$a_{11} = 1$	$a_{12} = 3$	$a_{21} = 2$	$a_{22} = 4$	$a_{31} = 3$	$a_{32} = 2$
$b_{111} = 2$	$b_{112} = 3$	$b_{121} = 3$	$b_{122} = 4$	$b_{211} = 2$	$b_{212} = 2$
$b_{221} = 1$	$b_{222} = 5$	$b_{311} = 4$	$b_{312} = 2$	$b_{321} = 3$	$b_{322} = 3$
$h_{111} = 5$	$h_{112} = 3$	$h_{121} = 6$	$h_{122} = 3$	$h_{131} = 5$	$h_{132} = 5$
$h_{211} = 5$	$h_{212} = 3$	$h_{221} = 6$	$h_{222} = 3$	$h_{231} = 5$	$h_{232} = 5$
$tr_{111} = 5$	$tr_{112} = 3$	$tr_{121} = 6$	$tr_{122} = 3$	$tr_{131} = 5$	$tr_{132} = 6$
$tr_{211} = 4$	$tr_{212} = 4$	$tr_{221} = 4$	$tr_{222} = 3$	$tr_{231} = 5$	$tr_{232} = 4$
$\tilde{a}_{11} = 0.1$	$\tilde{a}_{12} = 0.2$	$\tilde{a}_{21} = 0.2$	$\tilde{a}_{22} = 0.4$	$\tilde{a}_{31} = 0.3$	$\tilde{a}_{32} = 0.2$
$\tilde{b}_{111} = 0.2$	$\tilde{b}_{112} = 0.3$	$\tilde{b}_{121} = 0.3$	$\tilde{b}_{122} = 0.4$	$\tilde{b}_{211} = 0.2$	$\tilde{b}_{212} = 0.2$
$\tilde{b}_{221} = 0.1$	$\tilde{b}_{222} = 0.5$	$\tilde{b}_{311} = 0.4$	$\tilde{b}_{312} = 0.2$	$\tilde{b}_{321} = 0.3$	$\tilde{b}_{322} = 0.3$
$\tilde{h}_{111} = 0.5$	$\tilde{h}_{112} = 0.3$	$\tilde{h}_{121} = 0.6$	$\tilde{h}_{122} = 0.3$	$\tilde{h}_{131} = 0.5$	$\tilde{h}_{132} = 0.5$
$\tilde{h}_{211} = 0.5$	$\tilde{h}_{212} = 0.3$	$\tilde{h}_{221} = 0.6$	$\tilde{h}_{222} = 0.3$	$\tilde{h}_{231} = 0.5$	$\tilde{h}_{232} = 0.5$
$\widetilde{tr}_{111} = 0.5$	$\widetilde{tr}_{112} = 0.3$	$\widetilde{tr}_{121} = 0.6$	$\widetilde{tr}_{122} = 0.3$	$\widetilde{tr}_{131} = 0.5$	$\widetilde{tr}_{132} = 0.6$
$\tilde{tr}_{211} = 0.4$	$\tilde{tr}_{212} = 0.4$	$\tilde{tr}_{221} = 0.4$	$\tilde{tr}_{222} = 0.3$	$\tilde{tr}_{231} = 0.5$	$\tilde{tr}_{232} = 0.4$

$$\widehat{M}_{11} \sim \mathcal{N}(200,10),\ \ \widehat{M}_{12} \sim \mathcal{N}(300,20),\ \ \widehat{M}_{21} \sim \mathcal{N}(240,10),$$
$$\widehat{M}_{22} \sim \mathcal{N}(180,10),\ \ \widehat{M}_{31} \sim \mathcal{N}(280,20),\ \ \widehat{M}_{32} \sim \mathcal{N}(160,10).$$

The resources available to all the plants are 10000. The production capacities of the plants are 1800, 1500 and 1200, respectively. Assume that all $\theta_{kl} = 0.8$.

Let $C_P = C_D = 6000$. Suppose that the initial minimal satisfactory level as $\delta = 1.0$, and the lower and the upper bounds of Δ as $[0.6, 1.0]$. Using the algorithm described in Sect. 4, the decisions are obtained after three iterations:

$$\begin{aligned}
&X^*_{111} = 52.5,\ X^*_{112} = 22.1,\ X^*_{121} = 12.8,\ X^*_{122} = 26.8,\ X^*_{131} = 72.3,\ X^*_{132} = 44.2,\\
&X^*_{211} = 12.6,\ X^*_{212} = 62.5,\ X^*_{221} = 45.2,\ X^*_{222} = 92.5,\ X^*_{231} = 12.5,\ X^*_{232} = 14.5,\\
&Y^*_{111} = 82.5,\ Y^*_{112} = 72.5,\ Y^*_{121} = 62.5,\ Y^*_{122} = 20.2,\ Y^*_{211} = 32.0,\ Y^*_{212} = 26.5,\\
&Y^*_{221} = 65.3,\ Y^*_{222} = 34.5,\ Y^*_{311} = 17.5,\ Y^*_{312} = 42.5,\ Y^*_{321} = 54.3,\ Y^*_{322} = 62.5.
\end{aligned}$$

The corresponding costs of leader and follower are $Z^*_{DC} = 60512.8$ and $Z^*_{PC} = 32746.8$, respectively. The amounts of their emissions are $E^*DC = 26128.2$ and $E^*_{PC} = 11321.4$, respectively.

In order to test the effect the emission caps on costs, we use different emission caps to calculate the costs and actual emissions. For simplicity, let $C_{DC} = C_{PC}$. The results are shown by Fig. 2.

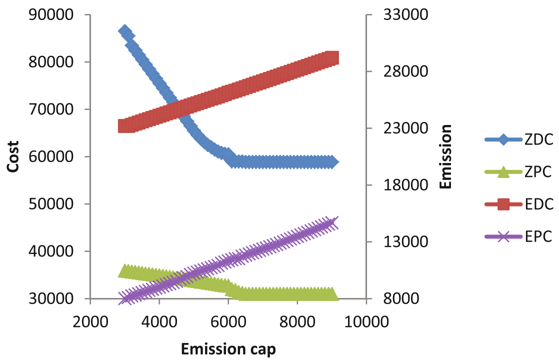


Fig. 2. The effect of emission cap on cost and carbon emission

Figure 2 shows the impact of varying the emission cap on the total cost and total emissions for the examples considered. Both the leader and follower have the similar trends. As expected, reducing the emission cap increases total cost and reduces total emissions. However, what is perhaps surprising is the fact that the emission cap can be significantly reduced without significantly affecting the total cost. In the example shown, reducing the emission cap from 6000 to 3000 reduces the average total amount of emissions by 50% but increases the average total cost by only 10%. Note that this reduction is obtained through adjustments in operational decisions only, namely, modifying order quantities in each period, could alone lead to significant reductions in carbon emissions while not significantly compromising overall cost.

6 Conclusions and Future Research

In this paper, we presented a bilevel decision-making model to illustrate how carbon footprint considerations could be incorporated into supply chain model. These insights highlight the impact of operational decisions on carbon emissions and the extent to which adjustments to operations can mitigate emissions. Our results show that operational adjustments alone can lead in some cases to significant emission reductions without significant increases in cost. More significantly, the results highlight the opportunity that exists to leverage collaboration across the supply chain to mitigate the cost of reducing emissions. Insights described in this paper revealed the importance of interactions between firms in determining overall emission levels and corresponding costs. The results in this paper will be helpful for the decision makers of supply chains to address challenges from strict emission caps regulation.

However, there are still some issues to worth attention.

(1) More quantitative analysis. The results in this paper are based on a numerical example. More quantitative analysis can reveal the essence behind the phenomenon, which is conducive to the application of the model.
(2) Scenario analysis under different regulatory emission control policies. In this paper, only the strict emission cap policy is considered. However, policies of the carbon tax, carbon offset and cap-and-trade are also used wildly. It is necessary to compare them deeply.

Acknowledgements. This research was supported by the Programs of NSFC (Grant No. 71401114), the Fundamental Research Funds for the Central Universities (Grant No. skqy201524).

References

1. Absi N, Dauzère-Pérès S et al (2013) Lot sizing with carbon emission constraints. Eur J Oper Res 227(1):55–61
2. Angelo JS, Barbosa HJ (2015) A study on the use of heuristics to solve a bilevel programming problem. Int Trans Oper Res 22(5):861–882
3. Baky IA (2014) Interactive topsis algorithms for solving multi-level non-linear multi-objective decision-making problems. Appl Math Model 38(4):1417–1433
4. Beheshti B, Özaltın OY et al (2015) Exact solution approach for a class of nonlinear bilevel knapsack problems. J Glob Optim 61(2):291–310
5. Benjaafar S, Li Y, Daskin M (2013) Carbon footprint and the management of supply chains: insights from simple models. IEEE Trans Autom Sci Eng 10(1):99–116
6. Bialas WF, Karwan MH (1984) Two-level linear programming. Manage Sci 30(8):1004–1020

7. Chang X, Xia H et al (2015) Production decisions in a hybrid manufacturing-remanufacturing system with carbon cap and trade mechanism. Int J Prod Econ 162:160–173
8. Charnes A, Cooper WW (1959) Chance-constrained programming. Manage Sci 6(1):73–79
9. Chen X, Benjaafar S, Elomri A (2013) The carbon-constrained eoq. Oper Res Lett 41(2):172–179
10. He P, Zhang W et al (2015) Production lot-sizing and carbon emissions under cap-and-trade and carbon tax regulations. J Cleaner Prod 103:241–248
11. Helmrich MJR, Jans R et al (2015) The economic lot-sizing problem with an emission capacity constraint. Eur J Oper Res 241(1):50–62
12. Hsueh CF (2015) A bilevel programming model for corporate social responsibility collaboration in sustainable supply chain management. Transp Res Part E Logistics Transp Rev 73:84–95
13. Hua G, Cheng T, Wang S (2011) Managing carbon footprints in inventory management. Int J Prod Econ 132(2):178–185
14. Kuo R, Lee Y et al (2015) Solving bi-level linear programming problem through hybrid of immune genetic algorithm and particle swarm optimization algorithm. Appl Math Comput 266:1013–1026
15. Lee KH (2011) Integrating carbon footprint into supply chain management: the case of hyundai motor company (hmc) in the automobile industry. J Cleaner Prod 19(11):1216–1223
16. Lv Y, Wan Z (2015) Solving linear bilevel multiobjective programming problem via exact penalty function approach. J Inequalities Appl 1:1–12
17. Ma W, Wang M, Zhu X (2014) Improved particle swarm optimization based approach for bilevel programming problem-an application on supply chain model. Int J Mach Learn Cybern 5(2):281–292
18. Rahmani A, MirHassani S (2015) Lagrangean relaxation-based algorithm for bi-level problems. Optim Methods Softw 30(1):1–14
19. Sadigh AN, Mozafari M, Karimi B (2012) Manufacturer-retailer supply chain coordination: a bi-level programming approach. Adv Eng Softw 45(1):144–152
20. Sakawa M, Matsui T (2014) Interactive fuzzy stochastic multi-level 0–1 programming using tabu search and probability maximization. Expert Syst Appl 41(6):2957–2963
21. Sakawa M, Nishizaki I, Uemura Y (1998) Interactive fuzzy programming for multilevel linear programming problems. Comput Math Appl 36(2):71–86
22. Wang D, Du G et al (2016) A stackelberg game theoretic model for optimizing product family architecting with supply chain consideration. Int J Prod Econ 172:1–18
23. Yeh K, Whittaker C et al (2015) Two stage stochastic bilevel programming model of a pre-established timberlands supply chain with biorefinery investment interests. Comput Chem Eng 73:141–153
24. Zheng Y, Liu J, Wan Z (2014) Interactive fuzzy decision making method for solving bilevel programming problem. Appl Math Model 38(13):3136–3141

Supply Chain Coordination by Revenue Sharing Contract Under Different Carbon Emission Policies

Li Lu(✉)

Tourism School, Sichuan University, Chengdu 610065, People's Republic of China
luli_rudy@126.com

Abstract. To solve the problem of global warming, governments have carried out some corresponding policies to reduce carbon emissions. In this paper, we considered a two-echelon supply chain consisting of a supplier and a manufacturer, and researched the supply chain coordination by revenue sharing contract under two different carbon emission policies, Mandatory carbon emissions capacity and Cap-and-trade. The study indicates that: under both two carbon policies, the revenue sharing contract can coordinate the supply chain. Manufacturer's revenue sharing proportion under cap-and-trade is always less than the situation of no carbon emissions constraint. The expected profits of manufacturer under cap-and-trade can be greater than the situation without carbon emissions constraints in certain conditions.

Keywords: Supply chain coordination · Revenue sharing contract · Mandatory carbon emissions capacity · Cap-and-trade

1 Introduction

It is a common knowledge that large amount of carbon emissions directly leads to global warming which has brought serious challenges to human survival and development. The best way to solve this problem is changing the way of production and life. So realize the sustainable development of low-carbon economy is becoming a hot spot of global attention. Nowadays, different governments including China have carried out the corresponding carbon emissions regulation policy according to their national conditions. Typical carbon policies are Mandatory carbon emissions capacity (Cap) and Cap-and-trade. Both the two policies are carbon constraints under the regulation of government, and the government will assign different initial carbon emissions quota to different enterprises. As we know, cap policy is a kind of mandatory constraint. The enterprises cannot sale or buy the carbon emissions quota in carbon trading market when the initial is overmuch or insufficient. However, cap-and-trade policy allows enterprises sale or buy the carbon emissions quota in carbon trading market when the initial is overmuch or insufficient, in order to meet the production requirements and

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_89

improve yields [13]. The earliest implementation of Cap and Cap-and-trade policies is European Emissions Trading System (EU ETS), and has now become the world's largest carbon emissions trading market [2].

The implementation of carbon policy in operation management has brought new challenges to the enterprise. It makes enterprise management more complex in performance target for decision-making (dual goals of increasing profits and reducing carbon emissions), the decision variables (production, order, pricing and other traditional decision variables and carbon emissions, carbon trading decision variables) and decision environment (production capacity, capital and other traditional and limitation of carbon constraints). As we all know, carbon emissions are throughout the entire supply chain, however, the existing researches on carbon policies in practice are mainly based on the perspective of individual enterprise. If only focus on individual enterprise, and cannot effectively solve the supply chain coordination problem of upstream and downstream enterprises, we cannot fundamentally promote carbon policy's implementation, and achieve the ultimate goal of reducing carbon emission [15].

Revenue sharing contract is a common supply chain coordination mechanism. Downstream enterprise buys goods from upstream enterprise at a relatively low wholesale price, but shares the sales revenue with upstream enterprise [3]. In this paper, we analyze the supply chain coordination problem based on revenue sharing contract, and research to solve the following problems:

(1) Whether revenue sharing contract could coordinate the supply chain under different carbon policy?
(2) How does the different carbon policy affect the revenue sharing ratio?
(3) How does the different carbon policy affect the manufacturer's expected profit?

The rest of this paper is organized as follows: Sect. 2 reviews the relevant literatures; Sect. 3 describes the problem and assumptions; Sect. 4 presents the model and analysis; and discussion is presented in Sect. 5; finally Sect. 6 concludes this paper.

2 Literature Review

In this paper, we review literature spanning across two streams. The first section addresses literature on supply chain coordination by revenue sharing contract. The second section addresses literature on supply chain coordination under different carbon policies.

2.1 Supply Chain Coordination by Revenue Sharing Contract

From the perspective of economics, Mortimer et al. [16] carried on the empirical research on the application of revenue sharing contract in the DVD rental industry. Giannoccaro et al. [10] systematically studied the problem of supply chain coordination by revenue sharing contract. Cachon and Lariviere [4] comprehensively analyzed of the advantages and limitations of revenue sharing contract. The results showed, although the revenue sharing contract could achieve supply chain coordination, the upstream enterprise did not achieve the optimal as this mechanism was instable. Kong et al. [14] analyzed the supply chain coordination by revenue sharing contract under the circumstance of information disclosure. Govindan et al. [11] took PC industry as an example, and analyzed the reverse supply chain coordination problem based on revenue sharing contract. The results indicated that revenue sharing contract could coordinate the reverse supply chain effectively. Arani et al. [1] introduced a novel mixed revenue-sharing option contract to coordinate a retailerCmanufacturer supply chain.

2.2 Supply Chain Coordination Under Different Carbon Policies

Ghosh et al. [8] took garment industry as an example, and studied the influence of green level on prices and profits, then put forward a kind of contract including two parts of tax to coordinate the green supply chain. Swami et al. [17] considered the consumers' environmental consciousness, and thought that green operation efforts of the retailers and manufacturers in the supply chain would affect consumer demand, then designed a cost sharing contract to realize supply chain coordination. Jaber et al. [12] considered a two-stage supply chain consisting a manufacturer and a retailer, and studied the supply chain coordination mechanism under Cap-and-trade policy when manufacturer undertook the cost of carbon emission. Based on the newsboy model, Choi [6] analyzes the impact of wholesale price and price subsidies contract under Cap-and-trade policy on retailers procurement source selection. Du et al. [7] considered a two-stage supply chain consisting of a carbon emission dependent enterprise and a carbon emission supplier under Cap-and-trade policy, and designed a supply chain coordination mechanism using the non-cooperative game theory. Ghosh and Shah [9] explored supply chain coordination issues arising out of green supply chain initiatives and explore the impact of cost sharing contract on the key decisions of supply chain players undertaking green initiatives. Xu [18] analyzed the decision behavior and coordination mechanisms for a two-echelon sustainable supply chain under a cap-and-trade regulation. Cao et al. [5] investigated the government's role in allocating the appropriate emission quota to maximize social members' utilities and analyzes how the emission-dependent enterprise improves revenues of both itself and the whole system through supply chain collaboration.

Through literature review, we found that the revenue sharing contract was a widely used supply chain coordination mechanism. But most of the researches did not consider the impact of different carbon policy on supply chain coordination mechanism. In order to fill the gap, we will analyze the supply chain coordination problem based on revenue sharing contract under different carbon policies.

3 Problem Description and Assumption

In this paper, taking the two echelon supply chain (one supplier and one manufacturer) which is dominant by supplier as the research object, we analyze the supply chain coordination problem based on revenue sharing contract under Cap and Cap-and-trade policy.

Before the sales season, the manufacturer purchases raw materials Q at wholesale price w from supplier. The manufacturer faces the random demand D with probability density $f(x)$ and distribution function $F(x)$, and sells the product at retail price p. At the end of the sales season, the salvage value per unit product is v, the shortage cost per unit product is g. The manufacture keeps $\varphi_i \in [0,1]$ $(i = 1,2,3)$ part of own revenue, and gives $1-\varphi_i$ part to the supplier, where $i = 1,2,3$ respectively represents the situation of no carbon constraint, Cap policy and Cap-and-trade policy. The supplier's purchasing cost is c_{s}, the manufacture's producing cost is c_{m}, carbon emission for producing unit product is e, trading price for unit carbon emission quota is k, and the carbon Cap set by the government is E.

For the convenience of description, considering the decentralized decision, let $Q_{\varphi 1}^*$, $Q_{\varphi 2}^*$, $Q_{\varphi 3}^*$ present manufacture's optimal order quantity under no carbon constraint, Cap and Cap-and-trade policy respectively; considering the integrated decision, let Q_1^*, Q_2^* and Q_3^* present supply chain's optimal order quantity under no carbon constraint, Cap and Cap-and-trade policy respectively. In addition, for the convenience of research, we make the following assumptions:

(1) The supplier and the manufacturer are risk neutral and perfectly rational, and the information between supply chain enterprises is completely symmetrical.
(2) The product's selling cycle is short, the market is stable, and during the sales period, the product's price remains the same.
(3) Carbon policy only applies to manufacturer.
(4) $p > w + c_m$ indicates that the sale price of product is always greater than the manufacturer's production costs, and the manufacturer is profitable.
(5) $w > c_s$ indicates that supplier's wholesale price is always greater than cost, supplier is also profitable.

4 Model and Analysis

4.1 The Base Model Without Carbon Emission Constraint

In this scenario, we get the manufacturer's expected profit

$$E\left[\pi_{m1}\left(Q_{\varphi 1}\right)\right] = \varphi_1 p S\left(Q_{\varphi 1}\right) + \varphi_1 v I\left(Q_{\varphi 1}\right) - g L\left(Q_{\varphi 1}\right) - \left(w + c_m\right) Q_{\varphi 1}, \quad (1)$$

where $S\left(Q_{\varphi 1}\right) = Q - \int_0^{Q_{\varphi 1}} f\left(x\right) dx$ presents the expected sales volume, $I\left(Q_{\varphi 1}\right) = \int_0^{Q_{\varphi 1}} \left(Q_{\varphi 1} - x\right) f\left(x\right) dx$ presents the expected inventory volume, and $L\left(Q_{\varphi 1}\right) = \int_{Q_{\varphi 1}}^{+\infty} \left(x - Q_{\varphi 1}\right) f\left(x\right) dx$ presents the expected shortage volume.

While the supplier's expected profit is

$$E\left[\pi_{s1}\left(Q_{\varphi1}\right)\right]=\left[w-c_s+\left(1-\varphi_1\right)v\right]Q_{\varphi1}+\left(1-\varphi_1\right)\left(p-v\right)S\left(Q_{\varphi1}\right). \qquad (2)$$

As $\frac{dE[\pi_{m1}(Q_{\varphi1})]}{dQ_{\varphi1}}=\left[\varphi_1\left(p-v\right)+g\right]\left[1-F\left(Q_{\varphi1}\right)\right]-\left(w+c_m-\varphi_1 v\right)$, $\frac{d^2E[\pi_{m1}(Q_{\varphi1})]}{d^2Q_{\varphi1}}=-\left[\varphi_1\left(p-v\right)+g\right]f\left(Q_{\varphi1}\right)<0$, there exists a unique optimal order quantity $Q^*_{\varphi1}$ which maximizes $E\left[\pi_{m1}\left(Q_{\varphi1}\right)\right]$.

Let $\frac{dE[\pi_{m1}(Q_{\varphi1})]}{dQ_{\varphi1}}=0$, we get optimal order quantity $Q^*_{\varphi1}=F^{-1}\left(\frac{\varphi_1 p+g-w-c_m}{\varphi_1(p-v)+g}\right)$. To coordinate the supply chain, we have to make $Q^*_1=Q^*_{\varphi1}$, that is $F^{-1}\left(\frac{p+g-c_s-c_m}{p+g-v}\right)=F^{-1}\left(\frac{\varphi_1 p+g-w-c_m}{\varphi_1(p-v)+g}\right)$, we get $\frac{p+g-c_s-c_m}{p-v+g}=\frac{\varphi_1 p+g-w-c_m}{\varphi_1(p-v)+g}$, so $\varphi_1=\frac{(p-v)(c_m+w)+(w-c_s)g+vg}{(p-v)(c_m+c_s)+vg}$.

In order to guarantee both sides accept this contract, we must ensure that their respective benefits under the revenue sharing contract are greater than those before, namely realizing Pareto improvement. We should have

$$\frac{E\left[\pi_{s1}\left(Q^*_1\right)\right]}{E\left[\pi_{s1}\left(Q^*_{\varphi1}\right)\right]}>1, \qquad (3)$$

$$\frac{E\left[\pi_{m1}\left(Q^*_1\right)\right]}{E\left[\pi_{m1}\left(Q^*_{\varphi1}\right)\right]}>1. \qquad (4)$$

As $E\left[\pi_{s1}\left(Q^*_1\right)\right]+E\left[\pi_{m1}\left(Q^*_1\right)\right]=E\left[\pi_{T1}\left(Q^*_1\right)\right]$, $E\left[\pi_{s1}\left(Q^*_1\right)\right]=\left(1-\varphi_1\right)E\left[\pi_{T1}\left(Q^*_1\right)\right]$, $E\left[\pi_{m1}\left(Q^*_1\right)\right]=\varphi_1 E\left[\pi_{T1}\left(Q^*_1\right)\right]$, put into the formula, we get $\frac{(1-\varphi_1)\,E[\pi_{T1}(Q^*_1)]}{E\left[\pi_{s1}\left(Q^*_{\varphi1}\right)\right]}>1$, $\frac{\varphi_1 E[\pi_{T1}(Q^*_1)]}{E\left[\pi_{m1}\left(Q^*_{\varphi1}\right)\right]}>1$. To simplify, we get $\frac{E\left[\pi_{m1}\left(Q^*_{\varphi1}\right)\right]}{E\left[\pi_{T1}\left(Q^*_1\right)\right]}<\varphi_1<1-\frac{E\left[\pi_{s1}\left(Q^*_{\varphi1}\right)\right]}{E\left[\pi_{T1}\left(Q^*_1\right)\right]}$.

The result indicates although revenue sharing contract can coordinate supply chain without carbon emission constraint, sharing proportion must meet a certain relationship to make the Pareto improvement for both sides; otherwise the contract shall be invalid.

According to the assumption, $\varphi_1\in[0,1]$, that is $0\le\frac{(p-v)(c_m+w)+(w-c_s)g+vg}{(p-v)(c_m+c_s)+vg}\le 1$, then $\left(p-v\right)\left(c_m+w\right)+\left(w-c_s\right)g\le\left(p-v\right)\left(c_m+c_s\right)$, we can proof $w\le c_s$.

It shows that supplier must make the wholesale price less than or equal to the purchase price to achieve supply chain coordination, then the supplier's profits drop. Therefore, the manufacturer must share enough revenue to make up for the loss of supplier, and make him gain more profits.

4.2 The Cap Model

In this scenario, we get the manufacturer's expected profit

$$\begin{cases} E\left[\pi_{m2}\left(Q_{\varphi2}\right)\right]=\varphi_2 pS\left(Q_{\varphi2}\right)+\varphi_2 vI\left(Q_{\varphi2}\right)-gL\left(Q_{\varphi2}\right)-\left(w+c_m\right)Q_{\varphi2} \\ s.t.\ \ Q_{\varphi2}\le\frac{E}{e} \end{cases} \qquad (5)$$

While the supplier's expected profit is

$$E\left[\pi_{s2}\left(Q_{\varphi 2}\right)\right]=\left[w-c_s+\left(1-\varphi_2\right)v\right]Q_{\varphi 2}+\left(1-\varphi_2\right)\left(p-v\right)S\left(Q_{\varphi 2}\right). \tag{6}$$

Proposition 1. *There exists a unique optimal order quantity which maximizes manufacturer's expected profit by revenue sharing contract under Cap policy. When $Q_{\varphi 2}^* \le \frac{E}{e}$, the optimal order quantity is $Q_{\varphi 2}^* = F^{-1}\left(\frac{\varphi_2 p+g-w-c_m}{\varphi_2(p-v)+g}\right)$, when $Q_{\varphi 2}^* > \frac{E}{e}$, it is $Q_{\varphi 2}^* = \frac{E}{e}$.*

Proof. As $\frac{dE[\pi_{m2}(Q_{\varphi 2})]}{dQ_{\varphi 2}} = \left[\varphi_2\left(p-v\right)+g\right]\left[1-F\left(Q_{\varphi 2}\right)\right]-\left(w+c_m-\varphi_2 v\right)$, then $\frac{d^2E[\pi_{m2}(Q_{\varphi 2})]}{d^2Q_{\varphi 2}} = -\left[\varphi_2\left(p-v\right)+g\right]f\left(Q_{\varphi 2}\right)<0$, there exists a unique optimal order quantity $Q_{\varphi 2}^*$ which maximizes $E\left[\pi_{m2}\left(Q_{\varphi 2}\right)\right]$.

Let $\frac{dE[\pi_{m2}(Q_{\varphi 2})]}{dQ_{\varphi 2}} = 0$, we get optimal order quantity $Q_{\varphi 2}^* = F^{-1}\left(\frac{\varphi_2 p+g-w-c_m}{\varphi_2(p-v)+g}\right)$.

When $Q_{\varphi 2}^* \le \frac{E}{e}$, the Cap policy doesn't play a role to the supply chain, it can be seen as the situation without carbon emission constraint, $Q_{\varphi 2}^* = F^{-1}\left(\frac{\varphi_2 p+g-w-c_m}{\varphi_2(p-v)+g}\right)$.

However, when $Q_{\varphi 2}^* > \frac{E}{e}$, the Cap policy plays a role to the supply chain, as the manufacturer's profit function is concave function, and it is monotone increasing at the left hand of $Q_{\varphi 2}^*$. That makes the optimal order quantity be $Q_{\varphi 2}^* = \frac{E}{e}$. This completes the proof.

□

Proposition 2. *When $Q_2^* \le \frac{E}{e}$, and $Q_2^* \le \frac{E}{e}$, the supply chain can coordinate by revenue sharing contract in the scenario of Cap policy where $\varphi_2 = \frac{(p-v)(c_m+w)+(w-c_s)g+vg}{(p-v)(c_m+c_s)+vg}$. And when $\varphi_2 \in \left(\frac{E\left[\pi_{m2}\left(Q_{\varphi 2}^*\right)\right]}{E\left[\pi_{T2}\left(Q_2^*\right)\right]}, 1-\frac{E\left[\pi_{s2}\left(Q_{\varphi 2}^*\right)\right]}{E\left[\pi_{T2}\left(Q_2^*\right)\right]}\right)$, it can realize Pareto improvement for both supplier and manufacturer; when $Q_{\varphi 2}^* \le \frac{E}{e} < Q_2^*$, the supply chain can coordinate where $\varphi_2 = \frac{g\left[1-F\left(\frac{E}{e}\right)\right]-w-c_m}{F\left(\frac{E}{e}\right)(p-v)-p}$, when $Q_2^* > Q_{\varphi 2}^* > \frac{E}{e}$, the supply chain can coordinate where $\varphi_2 = 1$.*

Proof. (1) When $Q_2^* \le \frac{E}{e}$ and $Q_{\varphi 2}^* \le \frac{E}{e}$, the Cap policy doesn't play a role. To achieve supply chain coordination, we should make $Q_2^* = Q_{\varphi 2}^*$, that is, $F^{-1}\left(\frac{p+g-c_s-c_m}{p+g-v}\right) = F^{-1}\left(\frac{\varphi_2 p+g-w-c_m}{\varphi_2(p-v)+g}\right)$, we get $\frac{p+g-c_s-c_m}{p-v+g} = \frac{\varphi_2 p+g-w-c_m}{\varphi_2(p-v)+g}$. To simplify, we obtain $\varphi_2 = \frac{(p-v)(c_m+w)+(w-c_s)g+vg}{(p-v)(c_m+c_s)+vg}$.

In order to make sure that both sides realize Pareto improvement, we should have $\frac{E[\pi_{s2}(Q_2^*)]}{E\left[\pi_{s2}\left(Q_{\varphi 2}^*\right)\right]} > 1$, $\frac{E[\pi_{m2}(Q_2^*)]}{E\left[\pi_{m2}\left(Q_{\varphi 2}^*\right)\right]} > 1$. As $E\left[\pi_{s2}\left(Q_2^*\right)\right] + E\left[\pi_{m2}\left(Q_2^*\right)\right] = E\left[\pi_{T2}\left(Q_2^*\right)\right]$, $E\left[\pi_{s2}\left(Q_2^*\right)\right] = \left(1-\varphi_2\right) E\left[\pi_{T2}\left(Q_2^*\right)\right]$, $E\left[\pi_{m2}\left(Q_2^*\right)\right] = \varphi_2 E\left[\pi_{T2}\left(Q_2^*\right)\right]$, put into the formula, we get $\frac{\varphi_2 E[\pi_{T2}(Q_2^*)]}{E\left[\pi_{m2}\left(Q_{\varphi 2}^*\right)\right]} > 1$. To simplify, we get $\frac{E\left[\pi_{m2}\left(Q_{\varphi 2}^*\right)\right]}{E\left[\pi_{T2}\left(Q_2^*\right)\right]} < \varphi_2 < 1-\frac{E\left[\pi_{s2}\left(Q_{\varphi 2}^*\right)\right]}{E\left[\pi_{T2}\left(Q_2^*\right)\right]}$.

(2) When $Q_{\varphi 2}^{*} \le \frac{E}{e} < Q_2^{*}$, the Cap policy plays a role. Let $Q_2^{*} = Q_{\varphi 2}^{*}$, that is $F^{-1}\left(\frac{\varphi_2 p+g-w-c_m}{\varphi_2(p-v)+g}\right) = \frac{E}{e}$. To simplify, we obtain $F\left(\frac{E}{e}\right) = \frac{\varphi_2 p+g-w-c_m}{\varphi_2(p-v)+g}$, then $\varphi_2 = \frac{g\left[1-F\left(\frac{E}{e}\right)\right]-w-c_m}{F\left(\frac{E}{e}\right)(p-v)-p}$.

(3) When $Q_2^{*} > Q_{\varphi 2}^{*} > \frac{E}{e}$, as $Q_2^{*} = Q_{\varphi 2}^{*} = \frac{E}{e}$, the supply chain doesn't need to coordinate, the manufacture keeps all the profit, now $\varphi_2 = 1$.

This completes the proof. □

This proposition suggests that, when carbon cap doesn't play a role, supply chain can coordinate by revenue sharing contract under Cap policy and realize Pareto improvement for both sides; when carbon cap plays a role, although we can find the condition which can achieve supply chain coordination, it is unable to realize the Pareto improvement for both sides. Because the Cap policy damages the whole supply chain's profit, it is difficult to achieve supply chain coordination in reality.

4.3 The Cap-and-Trade Model

In this scenario, we get the manufacturer's expected profit

$$\begin{cases} E\left[\pi_{m3}\left(Q_{\varphi 3}\right)\right] = \varphi_3 pS\left(Q_{\varphi 3}\right) + \varphi_3 vI\left(Q_{\varphi 3}\right) - gL\left(Q_{\varphi 3}\right) & \text{when } Q_{\varphi 3}^{*} \ge \frac{E}{e} \\ \qquad - \left(w+c_m\right) Q_{\varphi 3} - k\left(eQ_{\varphi 3} - E\right) & \\ E\left[\pi_{m3}\left(Q_{\varphi 3}\right)\right] = \varphi_3 pS\left(Q_{\varphi 3}\right) + \varphi_3 vI\left(Q_{\varphi 3}\right) - gL\left(Q_{\varphi 3}\right) & \text{when} Q_{\varphi 3}^{*} < \frac{E}{e}. \\ \qquad - \left(w+c_m\right) Q_{\varphi 3} + \varphi_3 k\left(E - eQ_{\varphi 3}\right) & \end{cases} \tag{7}$$

When $eQ_{\varphi 3}^{*} \ge E$, $k\left(eQ_{\varphi 3} - E\right)$ presents the cost of buying carbon quota, and when $eQ_{\varphi 3}^{*} < E$, $k\left(E - eQ_{\varphi 3}\right)$ presents the revenue of selling carbon quota.

While the supplier's expected profit is

$$\begin{cases} E\left[\pi_{s3}\left(Q_{\varphi 3}\right)\right] = \left[w - c_s + \left(1-\varphi_3\right)v\right] Q_{\varphi 3} & \text{when } Q_{\varphi 3}^{*} \ge \frac{E}{e}; \\ \qquad + \left(1-\varphi_3\right)\left(p-v\right) S\left(Q_{\varphi 3}\right) & \\ E\left[\pi_{s3}\left(Q_{\varphi 3}\right)\right] = \left[w - c_s + \left(1-\varphi_3\right)v\right] Q_{\varphi 3} & \\ \qquad + \left(1-\varphi_3\right)\left[\left(p-v\right) S\left(Q_{\varphi 3}\right)\right. & \text{when } Q_{\varphi 3}^{*} < \frac{E}{e}. \\ \qquad \left. + k\left(E - eQ_{\varphi 3}\right)\right] & \end{cases} \tag{8}$$

Proposition 3. *There exists a unique optimal order quantity which maximizes manufacturer's expected profit by revenue sharing contract under Cap-and-trade policy. When* $Q_{\varphi 3}^{*} \ge \frac{E}{e}$ *the optimal order quantity is* $Q_{\varphi 3}^{*} = F^{-1}\left(\frac{\varphi_3 p+g-w-c_m-ke}{\varphi_3(p-v)+g}\right)$*; when* $Q_{\varphi 3}^{*} < \frac{E}{e}$*, it is* $Q_{\varphi 3}^{*} = F^{-1}\left(\frac{\varphi_3 p+g-w-c_m-\varphi_3 ke}{\varphi_3(p-v)+g}\right)$.

Proof. (1) When $Q_{\varphi 3}^{*} \ge \frac{E}{e}$, $E\left[\pi_{m3}\left(Q_{\varphi 3}\right)\right] = \left[\varphi_3\left(p-v\right)+g\right] S\left(Q_{\varphi 3}\right) - \left(w + c_m + ke - \varphi_3 v\right) Q_{\varphi 3} - gu + kE$. As $\frac{dE\left[\pi_{m3}\left(Q_{\varphi 3}\right)\right]}{dQ_{\varphi 3}} = \left[\varphi_3\left(p-v\right)+g\right]\left[1 - F\left(Q_{\varphi 3}\right)\right] - \left(w + c_m + ke - \varphi_3 v\right)$, $\frac{d^2 E\left[\pi_{m3}\left(Q_{\varphi 3}\right)\right]}{d^2 Q_{\varphi 3}} = -\left[\varphi_3\left(p-v\right)+g\right] f\left(Q_{\varphi 3}\right) < 0$, there exists a unique optimal order quantity $Q_{\varphi 3}^{*}$ which maximizes $E\left[\pi_{m3}\left(Q_{\varphi 3}\right)\right]$.

Let $\frac{dE[\pi_{m3}(Q_{\varphi 3})]}{dQ_{\varphi 3}} = 0$, we get optimal order quantity $Q^*_{\varphi 3} = F^{-1}\left(\frac{\varphi_3 p+g-w-c_m-ke}{\varphi_3(p-v)+g}\right)$.

(2) When $Q^*_{\varphi 3} < \frac{E}{e}$, $E\left[\pi_{m3}\left(Q_{\varphi 3}\right)\right] = \left[\varphi_3\left(p-v\right)+g\right]S\left(Q_{\varphi 3}\right) - \left(w+c_m+\varphi_3 ke-\varphi_3 v\right)Q_{\varphi 3} - gu + \varphi_3 kE$. As $\frac{dE[\pi_{m3}(Q_{\varphi 3})]}{dQ_{\varphi 3}} = \left[\varphi_3\left(p-v\right)+g\right]\left[1-F\left(Q_{\varphi 3}\right)\right] - \left(w+c_m + \varphi_3 ke-\varphi_3 v\right)$, $\frac{d^2E[\pi_{m3}(Q_{\varphi 3})]}{d^2Q_{\varphi 3}} = -\left[\varphi_3\left(p-v\right)+g\right]f\left(Q_{\varphi 3}\right) < 0$, there also exists a unique optimal order quantity $Q^*_{\varphi 3}$ which maximizes $E\left[\pi_{m3}\left(Q_{\varphi 3}\right)\right]$.

Let $\frac{dE[\pi_{m3}(Q_{\varphi 3})]}{dQ_{\varphi 3}} = 0$, we get optimal order quantity $Q^*_{\varphi 3} = F^{-1}\left(\frac{\varphi_3 p+g-w-c_m-\varphi_3 ke}{\varphi_3(p-v)+g}\right)$. This completes the proof. □

Proposition 4. *When $Q^*_{\varphi 3} \geq \frac{E}{e}$, the supply chain can coordinate by revenue sharing contract in the scenario of Cap-and-trade policy where $\varphi_3 = \frac{(p-v)(c_m+w+ke)+(w-c_s)g+vg}{(p-v)(c_m+c_s+ke)+vg}$; when $Q^*_{\varphi 3} < \frac{E}{e}$, the supply chain can coordinate where $\varphi_3 = \frac{(p-v)(g-c_m-w)-(p-ke)g+(w-c_s)g}{(p-v)(g-c_m-c_s)-(p-ke)g}$. And when $\varphi_3 \in \left(\frac{E\left[\pi_{m3}\left(Q^*_{\varphi 3}\right)\right]}{E\left[\pi_{T3}\left(Q^*_3\right)\right]}, 1-\frac{E\left[\pi_{s3}\left(Q^*_{\varphi 3}\right)\right]}{E\left[\pi_{T3}\left(Q^*_3\right)\right]}\right)$, it can realize Pareto improvement for both supplier and manufacturer.*

Proof. (1) When $Q^*_{\varphi 3} \geq \frac{E}{e}$, let $Q^*_{\varphi 3} = Q^*_3$, that is $\frac{p+g-c_s-c_m-ke}{p-v+g} = \frac{\varphi_3 p+g-w-c_m-ke}{\varphi_3(p-v)+g}$, to simplify, we have $\varphi_3 = \frac{(p-v)(c_m+w+ke)+(w-c_s)g+vg}{(p-v)(c_m+c_s+ke)+vg}$.

(2) When $Q^*_{\varphi 3} < \frac{E}{e}$, let $Q^*_{\varphi 3} = Q^*_3$, that is $\frac{p+g-c_s-c_m-ke}{p-v+g} = \frac{\varphi_3 p+g-w-c_m-\varphi_3 ke}{\varphi_3(p-v)+g}$, to simplify, we obtain $\varphi_3 = \frac{(p-v)(g-c_m-w)-(p-ke)g+(w-c_s)g}{(p-v)(g-c_m-c_s)-(p-ke)g}$.

In order to make sure that both sides realize Pareto improvement, we should have $\frac{E[\pi_{s3}(Q^*_3)]}{E\left[\pi_{s3}\left(Q^*_{\varphi 3}\right)\right]} > 1$, $\frac{E[\pi_{m3}(Q^*_3)]}{E\left[\pi_{m3}\left(Q^*_{\varphi 3}\right)\right]} > 1$. As $E\left[\pi_{s3}\left(Q^*_3\right)\right] + E\left[\pi_{m3}\left(Q^*_3\right)\right] = E\left[\pi_{T3}\left(Q^*_3\right)\right]$, $E\left[\pi_{s3}\left(Q^*_3\right)\right] = \left(1-\varphi_3\right)E\left[\pi_{T3}\left(Q^*_3\right)\right]$, $E\left[\pi_{m3}\left(Q^*_3\right)\right] = \varphi_3 E\left[\pi_{T3}\left(Q^*_3\right)\right]$, put into the formula, we get $\frac{(1-\varphi_3)E[\pi_{T3}(Q^*_3)]}{E\left[\pi_{s3}\left(Q^*_{\varphi 3}\right)\right]} > 1$, $\frac{\varphi_3 E[\pi_{T3}(Q^*_3)]}{E\left[\pi_{m3}\left(Q^*_{\varphi 3}\right)\right]} > 1$. To simplify, we obtain $\frac{E\left[\pi_{m3}\left(Q^*_{\varphi 3}\right)\right]}{E\left[\pi_{T3}\left(Q^*_3\right)\right]} < \varphi_3 < 1-\frac{E\left[\pi_{s3}\left(Q^*_{\varphi 3}\right)\right]}{E\left[\pi_{T3}\left(Q^*_3\right)\right]}$.

This completes the proof. □

This proposition shows that, whether carbon emission quota is enough or not, supply chain can coordinate by revenue sharing contract under Cap-and-trade policy. However, sharing proportion must meet a certain relationship to make the Pareto improvement for both sides; otherwise the contract shall be invalid.

5 Discussions

Through the analysis above, we notice that revenue sharing contract can coordinate the supply chain under no carbon emission constraint, Cap and Cap-and-trade policy. Next, we will discuss the effect of carbon policy on revenue sharing ratio and manufacturer's expected profit.

5.1 Effect of Carbon Policy on Revenue Sharing Ratio

Proposition 5. *When* $eQ_2^* \leq E \leq EQ_3^*$*,* $\varphi_2 > \varphi_3$*.*

Proof. We have known when $Q_2^* \leq \frac{E}{e}$, $\varphi_2 = \frac{(p-v)(c_m+w)+(w-c_s)g+vg}{(p-v)(c_m+c_s)+vg}$, when $Q_3^* \geq \frac{E}{e}$, $\varphi_3 = \frac{(p-v)(c_m+w+ke)+(w-c_s)g+vg}{(p-v)(c_m+c_s+ke)+vg}$, $\varphi_2 - \varphi_3 = ke\,(p-v)\,(p-v+g)\,(c_s-w)$, according to the assumption $p > v$, we know $p - v > 0$, $p - v + g > 0$. So we mainly discuss whether $c_s - w$ is positive or negative.

As $\varphi_2 \in [0,1]$, that is $0 \leq \frac{(p-v)(c_m+w)+(w-c_s)g+vg}{(p-v)(c_m+c_s)+vg} \leq 1$, then $(p-v)\,(c_m+w) + (w-c_s)\,g \leq (p-v)\,(c_m+c_s)$, we can prove $w \leq c_s$. As $\varphi_3 \in [0,1]$, the same theory proves that $w \leq c_s$.

In conclusion, as $c_s - w > 0$, that is $\varphi_2 - \varphi_3 > 0$, then $\varphi_2 > \varphi_3$. This completes the proof. □

This proposition shows that, under cap policy, when the carbon cap doesn't play a role (that is no carbon emissions constraint), the manufacturer's revenue share proportion is greater than it under Cap-and-trade policy. Obviously, in the situation of no carbon emission constraint, manufacturer's profit is bigger, so he can share a higher proportion of revenue.

5.2 Effect of Carbon Policy on Manufacturer's Expected Profit

When the initial carbon quotas assigned by government are enough under cap policy, it seems to be a problem without carbon emission constraint. In this paper, taking the manufacturer's expected profit without carbon emission constraint as a benchmark, we want to study the effect of carbon policy on manufacturer's expected profit. In addition, as the supply chain is coordinated by revenue sharing contract, we adopt the whole supply chain profit to represent the manufacturer's expected profit.

Proposition 6. *In the situation* $Q_2^* \leq \frac{E}{e}$*, when* $E > E^*$*,* $E\left[\pi_{T3}\left(Q_3^*\right)\right] > E\left[\pi_{T2}\left(Q_2^*\right)\right]$*; when* $E = E^*$*,* $E\left[\pi_{T3}\left(Q_3^*\right)\right] = E\left[\pi_{T2}\left(Q_2^*\right)\right]$*; when* $E < E^*$*,* $E\left[\pi_{T3}\left(Q_3^*\right)\right] < E\left[\pi_{T2}\left(Q_2^*\right)\right]$*; where* $E^* = EQ_3^* + \frac{1}{k}\left[E\left[\pi_{T2}\left(Q_2^*\right)\right] - E\left[\pi_{T2}\left(Q_3^*\right)\right]\right]$*.*

Proof. Compare $E\left[\pi_T\left(Q_2\right)\right]$ and $E\left[\pi_T\left(Q_3\right)\right]$.

If $E \leq EQ_3^*$, $E\left[\pi_{T3}\left(Q_3^*\right)\right] = E\left[\pi_{T2}\left(Q_3^*\right)\right] - k\left(eQ_3^* - E\right) \leq E\left[\pi_{T2}\left(Q_3^*\right)\right] < E\left[\pi_{T2}\left(Q_2^*\right)\right]$, we can get $E\left[\pi_{T3}\left(Q_3^*\right)\right] < E\left[\pi_{T2}\left(Q_2^*\right)\right]$.

If $E \geq EQ_3^*$, $E\left[\pi_{T3}\left(Q_3^*\right)\right] = E\left[\pi_{T2}\left(Q_2^*\right)\right] + k\left(eQ_2^* - E\right) > E\left[\pi_{T2}\left(Q_2^*\right)\right]$, we can get $E\left[\pi_{T3}\left(Q_3^*\right)\right] > E\left[\pi_{T2}\left(Q_2^*\right)\right]$.

As a result, there always exists $E^* \in \left(eQ_3^*, eQ_2^*\right)$, satisfy $E\left[\pi_{T3}\left(Q_3^*\right)\right] = E\left[\pi_{T2}\left(Q_2^*\right)\right]$, that is $E\left[\pi_{T2}\left(Q_3^*\right)\right] - k\left(eQ_3^* - E\right) = E\left[\pi_{T2}\left(Q_2^*\right)\right]$.

So we get $E^* = EQ_3^* + \frac{1}{k}\left[E\left[\pi_{T2}\left(Q_2^*\right)\right] - E\left[\pi_{T2}\left(Q_3^*\right)\right]\right]$. This completes the proof. □

This proposition indicates that the manufacture's expected profit under Cap-and-trade policy is greater than it in the situation without carbon emissions constraint under certain conditions. This is because the Cap-and-trade policy is a kind of mixed strategy combining market operation and government regulation. It not only control the total emission amount, but also offer an opportunity for the enterprises to buy or sell the carbon emission quota, and make them to guarantee production or gain more profits.

6 Conclusion and Future Research

In this paper, taking the two echelon supply chain (one supplier and one manufacturer) which is dominant by supplier as the research object, we analyze the supply chain coordination problem based on revenue sharing contract under Cap and Cap-and-trade policy. It provides several interesting observations.

(1) Revenue sharing contract can coordinate the supply chain under both Cap and Cap-and-trade policy, and can realize the Pareto improvement for both sides under certain conditions.
(2) When the carbon policy plays a role, the manufacture's revenue share proportion under Cap-and-trade policy is always less than it without carbon emission constraint. The manufacturer would buy carbon emission quota for the limitation of it. It produced a certain cost and reduced profits, so he would distribute less profit to the supplier.
(3) Under certain conditions, the manufacture's expected profit under Cap-and-trade policy is greater than it in the situation without carbon emissions constraint. Because Cap-and-trade policy not only control the total emission amount, but also offer an opportunity for the enterprises to buy or sell the carbon emission quota, and make them to gain more profits. However, the Cap policy do not has such function. As a result, theoretically speaking, Cap-and-trade policy is better than Cap policy.

The conclusions of this paper is beneficial to the upstream and downstream enterprises in such supply chain, at the same time they can also provide theoretical guidance for carbon emissions policy for the government. In future study, we can extend our research from one-to-one two-stage supply chain to one-to-many or three-stage supply chain.

Acknowledgements. This research is supported by Fundamental Research Funds for the Central Universities (funded by Sichuan University, No. 2015SCU11072 and No. skzx2015-gb26), and Soft Science Project (No.2015-RK00-00244-ZF) funded by Chengdu Science and Technology Bureau.

References

1. Arani HV, Rabbani M, Rafiei H (2016) A revenue-sharing option contract toward coordination of supply chains. Int J Prod Econ 178:42–56
2. Böhringer C (2014) Two decades of european climate policy: a critical appraisal. Rev Environ Econ Policy 8(1):1–17
3. Cachon GP (2003) Supply chain coordination with contracts. Handbooks Oper Res Manag Sci 11:227–339 Supply Chain Management
4. Cachon GP, Lariviere MA (2005) Supply chain coordination with revenue-sharing contracts: strengths and limitations. Manag Sci 51(1):30–44
5. Cao J, Zhang X, Zhou G (2016) Supply chain coordination with revenue-sharing contracts considering carbon emissions and governmental policy making. Environ Prog Sustain Energy 2(35):479–488
6. Choi TM (2013) The international journal of advanced manufacturing technology. Int J Prod Res 1(6851):835–847
7. Du SF, Zhu LL, Liang L (2013) Emission-dependent supply chain and environment-policy-making in the cap-and-trade system. Energy Policy 57:61–67
8. Ghosh D, Shah J (2012) A comparative analysis of greening policies across supply chain structures. Int J Prod Econ 2(135):568–583
9. Ghosh D, Shah J (2014) Supply chain analysis under green sensitive consumer demand and cost sharing contract. Int J Prod Econ 164:319–329
10. Giannoccaro I, Pontrandolfo P (2004) Supply chain coordination by revenue sharing contracts. Int J Prod Econ 89(2):131–139
11. Govindan K, Popiuc MN (2014) Reverse supply chain coordination by revenue sharing contract: a case for the personal computers industry. Eur J Oper Res 2(223):326–336
12. Jaber MY, Glock CH, Saadany AME (2013) Supply chain coordination with emissions reduction incentives. Int J Prod Res 1(51):956–968
13. Keohane NO (2009) Cap and trade, rehabilitated: using tradable permits to control us greenhouse gases. Rev Environ Econ Policy 3(1):42–62
14. Kong G, Rajagopalan S, Zhang H (2013) Revenue sharing and information leakage in a supply chain. Manag Sci 59(3):556–572
15. Matthews HD, Gillett N et al (2009) The proportionality of global warming to cumulative carbon emissions. Nature 459:829–832
16. Mortimer JH (2002) The effects of revenue-sharing contracts on welfare in vertically separated markets: evidence from the video rent industry. University of Californian Los Angeles
17. Swami S, Shah J (2013) Channel coordination in green supply chain management. J Oper Res Soc 3(64):336–351
18. Xu J, Chen Y, Bai Q (2016) A two-echelon sustainable supply chain coordination under cap-and-trade regulation. J Cleaner Prod 135:42–56

Research on Integration of Livestock Products Supply Chain Based on the Optimal Match Between Supply and Demand

Liang Zhao, Yong Huang(✉), Zhusheng Liu, Mingcong Wu, and Lili Jiang

Business School, Sichuan University, Chengdu 610064, People's Republic of China
huangyong_scu@aliyun.com

Abstract. China is a big country on livestock production and Chinese livestock products output value is increased yearly. However, Chinese livestock products are still mainly dominated by self-employed farmer's production mode which leads to imbalance between supply and demand, periodic fluctuations in prices. In this paper an integration of livestock products supply chain is proposed to solve the imbalance problem. A sample of a duck farm is chosen to show the process of the integration. First a discrete-event simulation model is built by the supply chain which is composed of hatching, breeding, and packaging and transporting. Then production plan, transport routes and the number of fixed assets are adjusted to achieve the goal that the mean of supply-demand ratio is close to 1 and the variance is close to 0.

Keywords: Livestock products · Supply chain · Duck farm · Integration · Discrete-event simulation

1 Introduction

China is a big country on livestock production and Chinese livestock products output value is increased yearly. However, the construction of livestock products supply chain (LPSC) is inefficient. At present, Chinese big enterprises in the livestock products market have just started the integration of supply chain, and the livestock products market is still mainly dominated by self-employed farmer's production mode. Farmers neither share their information nor have any plans for production, so it leads to the imbalance between production and market demand. Once a product is sold well, farmers flock to produce it blindly. Then the price will fall, and farmers will switch to produce other products sold well, finally becoming a vicious circle. In addition, the production of livestock products has many shortcomings, such as low levels of automation, decentralized farmers. If we want to integrate the supply chain, the issue that matches supply with demand must be met. When the farmers make production plans, they're faced with the situation that demand is stochastic and production plans must be decided before they know the real demand [1]. From the experience of developed countries

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_90

for agricultural development, scale management on livestock products has an incomparable cost advantage. Therefore, this paper considers building integrated livestock products supply chain (ILPSC) to solve the mismatch between supply and demand.

The process of production, cultivation, processing, transportation and sale in LPSC will produce added value. Therefore, the LPSC is a value-added supply chain from upstream to downstream. However, if a peasant-oriented platform is just simply built, the information provided for each other may not be entirely true due to the opportunism between enterprises in supply chain. It is difficult to realize the information sharing in practice and this will cause supply and demand imbalance, eventually to make market price fluctuations. Building the ILPSC and formation of close cooperation are the emphases, which can integrate the information flow in the supply chain. But in the implementation process, there are still many issues to be resolved, such as the determination of the assets number, production planning and transportation routes. This paper builds an ILPSC model with simulation software, aimed at matching supply with demand, and eventually using a sample to prove the effectiveness and feasibility of the ILPSC model.

This paper is organized as follows. In Sect. 2, relevant literatures are reviewed. In Sect. 3 the ILPSC model is shown and analyzed. In Sect. 4, an empirical model of duck supply chain is presented and the supply-demand radio of the model is optimized. Section 5 gives some conclusions.

2 Literature Review

LPSC is a kind of agricultural products supply chain. Chinese research on agricultural supply chain started in 1999, and LPSC was later introduced. Chen and Luo [2] put forward the idea of building a new Chinese LPSC by adding the information agent to provide a scientific decision-making for strengthening the supply chain management of Chinese LPSC. Zhou [15] proposed some conclusions to promote the healthy and stable development of the pig industry. Yu and Li [14] proposed the market operation mode and structure of livestock products were no longer a single entity competition of production, processing and sales. The survival and development of livestock products were more and more dependent on the overall quality and comprehensive strength of the whole industry chain. In the basis of this, the supply chain model of the combination of third-party logistics and livestock products was used to improve the whole efficiency of LPSC. Wang and Huang [12] put forward the SCOR model of agricultural products based on information network, and constructed the agricultural supply chain logistics management mode under the information network environment. Then through the empirical analysis of Shanghai freshwater supply chain logistics management mode, they proved the effectiveness and practicability of agricultural logistics supply chain management mode. Liu, Sun and Wang [4] found that the researches of Chinese agricultural supply chain were mainly concentrated on the traditional mode of operation of agricultural supply chain, fresh agricultural products at the end of the supply chain supermarket and information technology under the traceability system and e-commerce environment. And he believed that the establishment of a new integrated production and marketing supply chain and circulation system to achieve the circulation of agricultural products and information network management, could make Chinese agricultural products enter further into the global supply chain.

Compared with China, the United States, the Netherlands, the European Union, Japan and other developed countries have been perfect on the development of agricultural supply chain, so there are many researches about LPSC, especially about beef and pork supply chain. Birthal and Pratap [10] proposed that the revolutionary progress in livestock production was demand-driven. The income elasticity of demand was higher compared to most other food commodities. And livestock would have a larger effect on poverty reduction compared to crop sector. Facchioli [3] used a computational simulation system tool to coordinate its flow of matter and information. Piewthongngam [8] developed a system dynamics model as a tool for managers to visualise the movement of the entire production chain. William, Norina and Cassavant [13] thought SCRA was a “production adjustment & consumer-driven” systems. Van Roekel, Kopicki and Broekmans [11] gave the basic four steps in the development of agricultural supply chains. Simchi and Kaminsky [9] pointed out that information sharing and task planning was the key to integrating supply chain. Min and Zhou [5] summarized the methods and applications of integrated supply chain modeling. It provided a reference to establishment of the ILPSC. Pan and Jean [7] analyzed the link between pork supply chain in China and the United States, and put forward that a mechanism of enterprises and peasants, a sound logistics operation system and an information network platform should be established so that the efficiency of Chinese pork processing supply chain could be improved. Robert and Jon [6] used a visual modelling environment to overcome the problems involved in implementing simulation models.

3 Model Description

3.1 Integrated Livestock Products Supply Chain Model

The ILPSC model refers to the integration of all members in the LPSC, based on the common goals, to achieve true sharing of information and integration of supply, production and sale. Figure 1 shows the whole ILPSC model, and there are mainly four parts. The first part is the process of producers which contains capital goods suppliers and flows. Capital goods suppliers weekly produce pups as per production plan which is designed by demand. Then, a collection of several suppliers makes up a flow. The combined production of a flow is kept together when it moves through the model. The second part is the process of manufacturers for a period of breeding activities. The third part is the process of processors where livestock products are processed and packaged. The fourth part is the process of sending products to market.

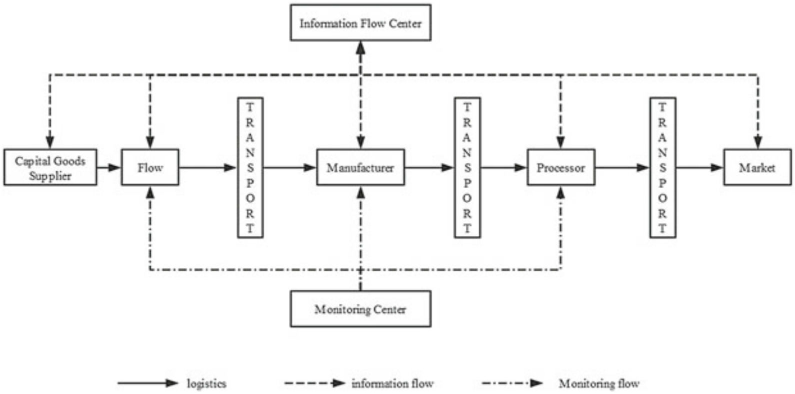


Fig. 1. The integrated livestock products supply chain model

In the whole model, the information flow center makes information flow in the whole supply chain smoothly and shares the information with every part in the model. At the same time, reasonable decisions are made on all links of the supply chain through the integrated information, such as production plan and transportation routes. And early in the supply chain construction, it can determine whether the number of assets to meet the market demand number by the weekly market demand plan. Monitoring center is mainly for flows, manufacturers and processors. Among them, the flows and manufacturers are mainly concerned about the health condition of livestock. What we want to monitor is whether the livestock are sick, whether the disease is contagious and whether the sick livestock are dead. In addition, the weekly sales of processors should be monitored, because the processor needs to sell products to the market weekly. And a batch of products is sent to market several weeks.

The ILPSC model is a two-way driving model. That is, on the one hand, the model needs the production plan of the most upstream companies to push the entire supply chain, on the other hand, it needs the final market demand to pull. Therefore, it's necessary to forecast market demand to pull the production to guarantee the matching of supply and demand information, and then push the production through the production plan. Events in the model take 1 week.

The goal of the model is to strike a balance between the supply of livestock products to market and the demand of the market. The supply-demand ratio can be used to response the situation about balance of supply and demand. Therefore, whether the target is achieved or not can be determined by analyzing the mean and variance of all processors' corresponded supply-demand ratio. If the mean of it is close to 1 and the variance is close to 0, it can be proved that the target is achieved. And a sub-goal may exist, such as the cost. The rule to transport products can be considered, when the transporter works. Distance between different parts in the model is affected by the rule, and the cost is affected by the distance. There are some uncertain factors in the model which is concerned with the goal of the model. The first one is the uncertainties associated with the health condition of livestock. It includes whether the livestock are diseased, which type of health episodes is caught, whether it has died after the illness and which week after the illness is the time livestock dead. Secondly, the week when the livestock products are sent to the market is also an uncertainty factor. All of them will affect the number transported to the market, through the uncertain production losses for the entire system caused by themselves. In this paper, these uncertainties are solved by simulation model.

The ILPSC model is suitable for livestock products, which mainly refer to livestock and poultry meat, and livestock including pigs, cattle, sheep and so on, poultry including chickens, ducks, geese and other poultry meat and wild birds. The characteristics of livestock products include that the number of production can be counted, production cycle is long, the sickness can result in uncertain production losses, it's easily restricted by the size of the farm, it will be transported between different farms and processing and packaging are needed.

3.2 Model Entities, Variables

To build a discrete-event simulation model, entities and variables must be determined.

Model entities. Livestock, transporters, capital goods suppliers, flows, manufacturers and processors are used in the ILPSC model as model entities.

Model variables. There are 12 variables in the ILPSC model, and the details are shown in Table 1.

3.3 Analysis of Model Process

There are many process in the model, and the processes of health condition and selection of flows to manufacturers are the most important and analyzed.

(1) Health Condition Process

Livestock may get sick during the breeding process, and the type of diseases may be different. The infectivity of different types of diseases is also different. If the disease is infectious, the diseased livestock need to be isolated from other healthy livestock. The flow chart of the process is shown in Fig. 2.

(2) Selection of Flows to Manufacturers Process (based on the shortest distance rule)

When livestock are transported from flows to manufacturers, a transportation rule can be considered such as the shortest distance rule. It is preferable to transport the products to the available manufacturer who is the nearest to the flow. At the same time, the number of livestock transported will be constrained by the capacity of transporters and manufacturer's assets. The flow chart of the process is shown in Fig. 3.

Table 1. Variables of ILPSC model

Variable name	Variable description	Notes
Transfer week	The number of weeks in the model before to be sent to the market	Updated weekly when the value is 0, it will be sent to the market in the week
Health condition	0 is showed healthy, 1 is showed sick	
Health episode	The type of illnesses, using 1 n as disease 1 disease n	
Sick week	The number of weeks before a sick livestock dead	Updated weekly when the value is 0, it will be dead due to illness
Week	The number of weeks the model runs	
Quarter	The number of quarters the model runs	
Min mile#	The nearest distance between the flow and the nearest available manufacturer	# is the number of flows, $1, 2, \cdots, m$
Supply#	The number of livestock products supplied to markets weekly	# is the number of processors, $1, 2, \cdots, t$
Transporter#	The number of livestock delivered by a transporter	# is the number of flows, $1, 2, \cdots, m$
Ill livestock#	The number of sick livestock per week for each flow	# is the number of flows, $1, 2, \cdots, m$
Rate#	Weekly supply-demand ratio	# is the number of processors, $1, 2, \cdots, t$
SD Rate#	The difference between the mean of the Rate# in the simulation period and 1 the value is closer to 0, the matching rate is higher	# is the number of processors, $1, 2, \cdots, t$

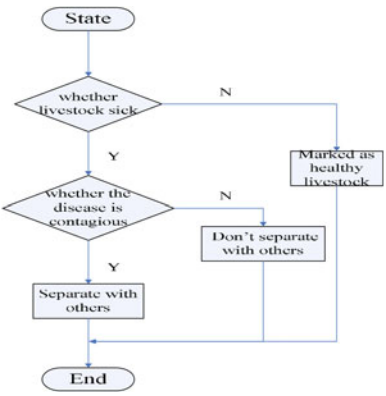


Fig. 2. The flow chart of the health condition process

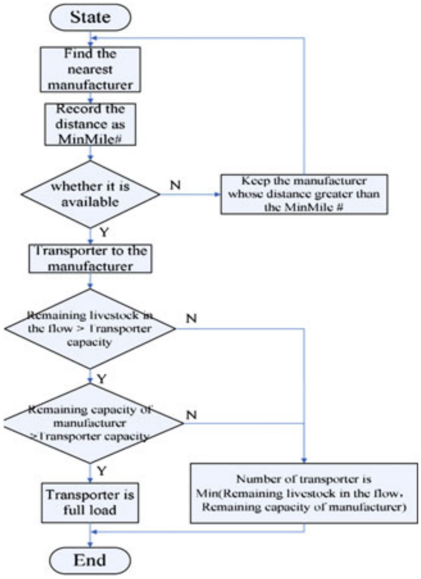


Fig. 3. The flow chart of the selection of flows to manufacturers process

3.4 Model Optimization Rules

To achieve the goal, the model needs to be optimized. There are many parts in the supply chain will have an impact on the goal. In this paper, three factors are mainly considered including the asset numbers of manufacturers, production plans and transportation routes.

(1) The asset numbers of manufacturers
It refers to the number of assets in the supply chain where manufacturers breed livestock. Manufacturers are used to breed pups provided by capital goods suppliers, and it's the key to connecting upstream and downstream of the supply chain. Therefore, when the production plans of capital goods suppliers have been decided with the change of the market demand, the asset number of manufacturers should change accordingly. So, whether the asset number of manufacturers is enough must be first determined in the establishment of the model. Only when the number of manufacturers' assets is sufficient can the model continue to be optimized.

(2) The production plans
A production plan is a schedule of the number of pups produced by capital goods suppliers. It directly determines the approximate size of the livestock products that are ultimately sent to the market. Therefore, the sum of livestock products transported to market will be closed to the sum of market demand by optimizing the production plan.

(3) The transportation routes
Transportation routes refer to the routes that transporters deliver livestock between the links of the supply chain. This paper mainly considers the route between manufacturers and processors. Since processors have their corresponding markets, processed livestock products can be distributed evenly by adjusting the routes, when the sum of supply is close to the demand.

4 Empirical Analysis

4.1 Model Overview

The empirical data in this article refers to the competition of the 2016 Arena Simulation Student Competition: Rockwell Duck Farm Supply Chain Optimization. Two schedules will drive the simulation model of the farm system: a market schedule, and a hatching farms schedule.

There are currently about 75 hatching farms in the system which produce a total of approximately 100,000 ducks each week. Hatching farms breed ducks per the hatching farms schedule, and a collection of several hatching farms makes up a flow to move together. The order for transport is carried out in descending order of flows and finds the nearest available growth/finish farms to receive the ducks with the shortest distance rules. Production will be affected by the health condition of ducks. There are four diseases to ducks showing in Table 2. The probability of getting sick in four quarters of every hatching farm and prevalence of the four diseases are based on the data provided in the competition.

Table 2. Four diseases to ducks

Disease types	Requires separation?	Sick weeks (week)	Output loss (%)
DED Exposed (Disease$_1$)	No	4	0.25
DED Naive (Disease$_2$)	No	4	1
DRRS Non-Vaccinated (Disease$_3$)	Yes	8	0.4
DRRS Vaccinated (Disease$_4$)	Yes	8	0.1

There are currently about 230 growth/finish farms in the system. The farms receive the ducks from the hatching farms until it is full. Ducks are breeding 22 weeks in the growth/finish farms. Because disease_3 and disease_4 need to be separated from other ducks, the growth/finish farm_27, growth/finish farm_199, growth/finish farm_1564, growth/finish farm_65, and growth/finish farm_179 are selected to receive the ducks suffering from the two diseases, which are relatively closer to each station.

There are 11 packers/plants in the system. After packaged and processed, ducks are sold for 6 weeks, the proportion shown in Table 3.

Table 3. Proportion of ducks sent to market weekly

Sales week	Sales ratio
Week 1	0.04
Week 2	0.08
Week 3	0.2
Week 4	0.25
Week 5	0.25
Week 6	0.18
Sum	1

Ducks are transported by truck in the system and each truck is loaded with 2,600 ducks. A distance matrix defines the miles between a central point of a flow's location to a particular growth/finish farm. The simulation time is two years or 104 weeks. Figure 4 shows the supply chain model of the duck farms established with visual simulation software. The lower left part of the facility interface is 75 hatching farms, the top part is 230 growth/finish farms and the lower right part is 11 packers/plants.

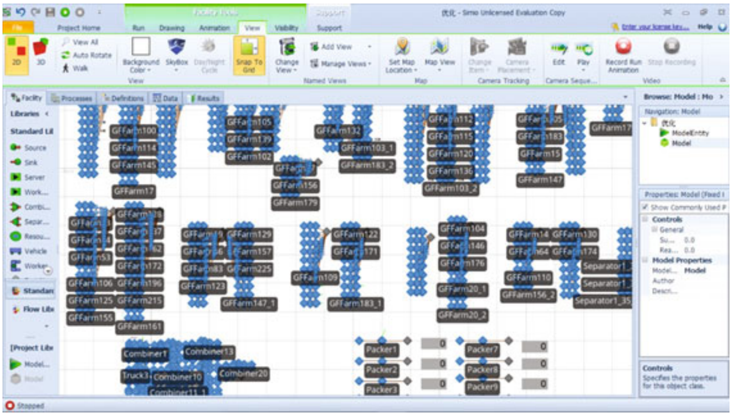


Fig. 4. The ILPSC model of duck farms established with visual simulation software

4.2 Model Optimization

When the model is run, it's found to be incorrect. The number of farms is insufficient to accommodate the duck. Therefore, there need more farms to meet the extra expectations. 16 growth/finish farms near the 16 flows are selected to form a group, three groups added. After running the model again, it's found that the supply of packer/plant$_5$ is zero. Therefore, transportation routes are adjusted per the results of operations. It's mainly solved through adjusting the corresponding packers/plants of the new three groups to the packers/plants whose supply is obviously insufficient. To find the optimal production plan, an experiment needs setting up. The controlled variable for the experiment is the number of hatchings per week for 75 hatching farms. The target is SD Rate# of 11 packers/plants. SD Rate# in 40 scenarios is shown in Fig. 5.

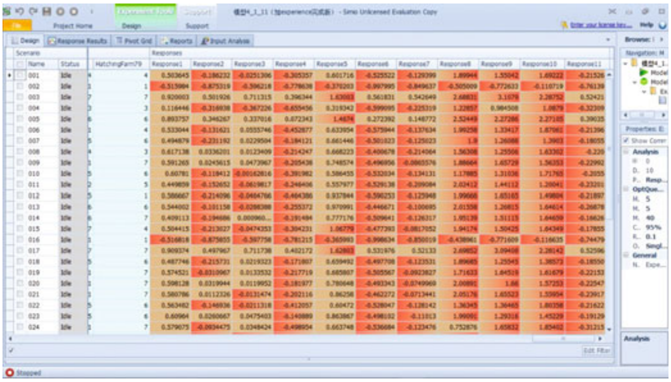


Fig. 5. SD Rate# in 40 scenarios after running experiment

Since the ultimate goal is to achieve the match of supply and demand, it can be achieved by finding a scenario where the supply-demand ratio is closer to 1 and the SD Rate# is closer to 0. Therefore, by comparing the mean and coefficient of variation (CV) of the absolute value of the SD Rate#, the scenario can be found where the mean and CV are closer to 0. It proves that the number of supply and demand is closer. An mean-CV line chart is shown in Fig. 6.

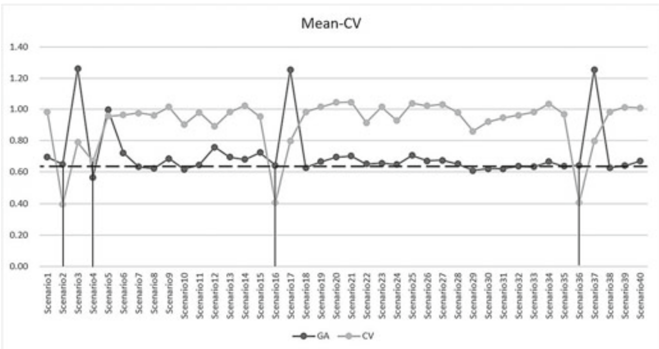


Fig. 6. Mean-CV line chart for 40 scenarios

Table 4. The adjustment scheme based on the result of experiment

Parker 1	Parker 2	Parker 3	Parker 4
G/F farm 15_4	G/F farm 65_4	G/F farm 63_2	G/F farm 54_4
G/F farm 20_4	G/F farm 179_4	G/F farm 20_2	G/F farm 183_4
G/F farm 63_4	G/F farm 224	G/F farm 36	G/F farm 16_4
G/F farm 147_4	G/F farm 178	G/F farm 207_4	G/F farm 76
	G/F farm 205_4	G/F farm 220_4	G/F farm 64
			G/F farm 16_2
			G/F farm 156_1
Parker 6	**Parker 7**	**Parker 8**	**Parker 10**
G/F farm 222	G/F farm 14	G/F farm 66	G/F farm 27_4
G/F farm 211	G/F farm 30	G/F farm 206	G/F farm 199_4
G/F farm 210	G/F farm 145		
G/F farm 209	G/F farm 208		**Parker 11**
G/F farm 95	G/F farm 34		G/F farm 156_4
G/F farm 213	G/F farm 17		
G/F farm 150	G/F farm 148		

The mean and CV of scenario 2,4,16,36 are all suitable. Finally, the production plan of scenario 4 is selected. Then the model is ran again and the transportation routes from the growth/finish farms to the packers/plants are adjusted based on the results. The adjustment scheme is shown in table 4(G/F farm represents Growth/Finish farm). Since the supplies of packer/plant$_5$ and packer/plant$_9$ are sufficient, they are not shown in Table 4.

The line graphs between the supply and demand of the 11 packaging plants in the original and the optimized model are shown in Fig. 7.

From Fig. 7, it can be seen the fluctuation range of the optimized model's supply is decreased compared with the original model, and supplies of every packer/plant are basically fluctuating up and down in demand. The supply-demand ratio of the original model and the optimized model is compared, as shown in Table 5. And the line chart is shown in Fig. 8.

According to Fig. 8, it can be seen that the overall supply-demand ratio of the 11 packers/plants of the optimized model is close to 1, and the fluctuation is obviously reduced. There are three changes seen from Table 5 the mean of supply-demand ratio of 11 packers/plants is reduced by 0.0836, the standard deviation is reduced by 0.6979 and the variation coefficient is reduced by 0.6478. Although the mean of the original model is also close to 1, the standard deviation is 0.8142 which means it is very volatile. And the standard deviation and the coefficient of variation after optimization is closer to 0, it can be seen that the optimized model can make the match of supply and demand better.

Fig. 7. The line graphs between the supply and demand of the 11 packaging plants in the original and the optimized model during 28-104th weeks

Table 5. The supply-demand ratio of the original model and the optimized model

Packer number	1	2	3	4	5	6	7
Supply-demand ratio	2.6161	1.0385	0.6777	1.1512	0	2.1518	0.6294
Optimized supply-demand ratio	0.994	1.0853	0.832	0.9447	1.0613	0.8967	0.7984
Packer number	8	9	10	11	Mean	*SD*	*CV*
Supply-demand ratio	0.4204	1.8445	0.7463	0.4056	1.062	0.8142	0.7667
Optimized supply-demand ratio	1.0332	1.1317	1.1111	0.8735	0.9784	0.1163	0.1189

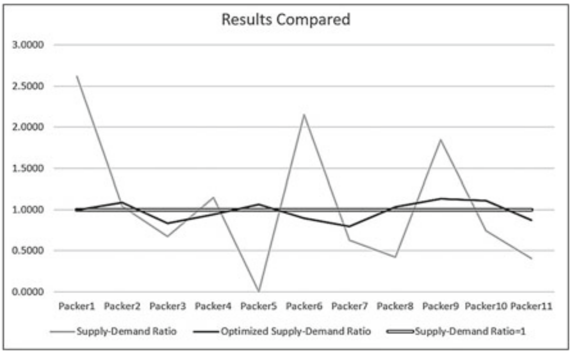


Fig. 8. The line chart for the supply-demand ratio of the original model and the optimized model

5 Conclusion

This paper establishes an ILPSC to integrate the entire supply chain, which ultimately matches the supply and demand of livestock products. The whole process of the ILPSC from the upstream capital goods suppliers to the downstream market is introduced, including the information flow center and the control center to control the information flows and uncertain factors. Then, an integrated supply chain model for a duck farm is constructed with simulation software. In the example, the number of assets in the growth farms is added, and the production plan of the hatching farms and the transportation routes between the growth farms and the packing plants are optimized. Finally, the mean of supply-demand ratio of the 11 packing plants is reduced by 0.0836, the standard deviation is reduced by 0.6979 and the variation coefficient is reduced by 0.6478. The construction of the ILPSC is a quite complex problem. Though the ILPSC can represent most of the integrated process of livestock products, it also can be extended to other aspects, for example, the question on reduction of transportation costs, and the question on increasing utilization of growth farms.

References

1. Cachon G, Terwiesch C (2009) Matching supply with demand, 2nd edn. McGraw-Hill, Singapore (in Chinese)
2. Chen C, Luo Y (2003) Construction on supply chain model of livestock products in China. J Nanjing Agric Univ 26:89–92 (in Chinese)
3. Facchioli P, Severino G et al (2015) Use of a simulation system tool at the logistic of a sugarcane company. Rev Metropolitana Sustentabilidade 5:112–127
4. Liu Z, Sun S, Wang J (2009) The trend of agricultural supply chain management in China. Commercial Res 3:161–164 (in Chinese)
5. Min H, Zhou G (2002) Supply chain modeling: past, present and future. Comput Ind Eng 43:231–249

6. Muetzelfeldt R, Massheder J (2003) The simile visual modelling environment. Eur J Agron Off J Eur Soc Agron 18:345–358
7. Pan C, Kinsey J (2002) The supply chain of pork: U.S. and China. Technical report, University of Minnesota
8. Piewthongngam K, Vijitnopparat P et al (2014) System dynamics modelling of an integrated pig production supply chain. Biosyst Eng 127:24–40
9. Simchi L, Kaminsky P, Simchi L (2000) Designing and managing the supply chain: concepts, strategies and case studies
10. Staal S (2015) Livestock marketing and supply chain management of livestock products. Indian J Agric Econ 70:89–92
11. Van R, Kopicki R et al (2002) Building agri supply chains: issues and guideline [eb/ol]
12. Wang N, Huang L (2005) A study on agricultural product logistics supply chain management mode based on information network. Res Agric Modernization 26:126–129 (in Chinese)
13. William C, Norina L, Cassavant K (2002) The use of supply chain management to increase exports of agricultural products
14. Yu X, Li C (2012) Research on supply chain of livestock products mode based on third-party logistics. In: Proceedings of 2011 International Conference on Engineering and Information Management (ICEIM 2011)
15. Zhou C (2016) An improved quantitative model for evaluating pork price fluctuation trend and the impaction in China: an empirical analysis between 2004–2014. Int J Future Gener Commun Netw 9:199–210 (in Chinese)

A Novel Multi-Objective Programming Model Based on Transportation Disruption in Supply Chain with Insurance Contract

Kai Kang(✉), Jing Zhao, and Yanfang Ma

School of Economics and Management, Hebei University of Technology,
Tianjin, China
kkang2000@sina.com

Abstract. Supply chain risk management is becoming increasingly a hot research issue. When the disruption occurs, insurance can be used to reduce the risk. In this paper, we set a novel multi-objective programming model based on transportation disruption in supply chain with insurance contract in order to protect the whole supply chain members. Then use a case to investigate the applicability of this model. From this mathematical model, we obtain the optimal order quantity and the pure premium, which can help the manufacture and the retailer make a better plan. And the insurance can protect the manufacturer and the retailer's profits. In the future, the insurance will be used in supply chains widely.

Keywords: Supply chain risk management · Transportation disruption · The optimal order quantity

1 Introduction

Since the 21st century, with the global procurement, non-core business outsourcing, and development of business models in supply chain management, such as the lean management, the space distance on the supply chain becomes longer and longer, while the time distance becomes shorter. The change of the temporal and spatial variation of the supply chain improves the possibility of an interrupt occurance. Every firm's supply chain is susceptible to a diverse set of risks, such as natural disasters, terrorism, war, financial crisis, supplier bankruptcy and transportation delays. A lot of strategies used by firms to mitigate disruption risks include emergence purchases, multi-sourcing, inventory reserve and some reliability improvement of supply process [6]. The awareness regarding the importance of supply chain risk management (SCRM) has grown in the recent years. Supply chains are becoming increasingly competitive and complex in order to effectively meet customer demands. Supply chain disruptions often lead to declining sales, cost increases, and service failures for the company. There are several kinds of mode of supply chain disruptions. Different scholars have different classifications for it. It can be clearly seen in Table 1.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_91

Table 1. The classifications of supply chain disruption

Year	Author	Classification
2013	Hishamuddin [3]	Supply disruption, transportation disruption
2015	Nooraie [7]	Financial risk, transportation disruption
2015	Heckmann [2]	Financial risk, demand disruption
2016	S. Tong [8]	Node disruption, chain disruption

Transportation disruption, in particular, is slightly different from other forms of SC disruptions, in that it only stops the flow of goods, whereas other disruptions may stop the production of goods as well. The transportation disruption of the supply chain is a non-negligible factor that it can reduce the manufacturer and the retailer's profits. For instance, in October 2000, two typhoons damaged the oil gas pipeline belonging to Shanghai Petroleum Natural Gas Co., Ltd. in China, leading to a 178-days service disruption of oil gas. Terrorist attacks on the world trade center on September 11, 2001, causing a stagnation of goods which came from the border of America. This led to a force to the Ford motor company to stop production activities temporarily. In March 2011, an earthquake happened in Japan, which caused 22 automobile manufacturing companies' shutdown, such as Toyota, Honda, nissan. Affected by the earthquake the Japanese exports of cars fell 12.5% in April compared with the year 2010, being the biggest drop since the 18 months. As you can see from the above cases, transportation disruption caused by emergency events not only leads manufacturers to stop production, but also brings irreparable damages to the downstream enterprises of the supply chain. Furthermore, a transportation disruption may affect the condition of the valuable goods in transit.

Once transportation interrupts, it has a great influence for each member on the supply chain, leading to transport delay, increasing the transportation time and the transportation cost, even bringing serious damage to related companies. Although realizing that there is a disruption risk in the supply chain, most of the enterprises in the aspect of risk management of supply chain disruptions have very little money and resources. To make sure for the firm's profit, we tend to make the insurance strategy according to the degree of risk. Purchasing an insurance for the emergency is a kind of protection for enterprises. The model that we have developed in this research addresses this vital aspect of transportation disruption.

This paper proposes a newly multi-objective programming model for a supply chain system subject to transportation disruption. In this model, we consider the insurance contract to deal with the transportation disruption risk.

The contents of this paper are organized as follows. Section 2 presents the related literature review. A description of the model and its formulation are given in Sect. 3. Section 4 shows the analysis of numerical example. Section 5 provides a conclusion of this paper and offers potential directions for future research.

2 Literature Review

There is a great deal of literature on supply chain. In recent years, the research in supply chain disruption is becoming a hot spot. Disruption has the characteristics of low probability of occurrence and big influence on the part of the supply chain, even the whole supply chain. Many scholars classified the types of supply chain disruption, and provided several operational strategies for managing disruption. But there are few papers which are mentioned the insurance to solve the risks.

Our study is related to studies focusing on the transportation disruption in supply chain. Zhen et al. [9] investigated four strategies: basic strategy, BI insurance strategy, backup transportation strategy, and mixed strategy, in order to deal with distribution centers' daily risk management. They used the mathematic model to compare BI insurance strategy with backup transportation strategy, and found that the choice of BI insurance strategy and the backup transportation strategy depended on transportation market, insurance market and distribution center's operational environments. Hishamuddin et al. [3] built a recovery model for a two-echelon serial supply chain when transportation disruption occurred. This model determined the optimal ordering and production quantities with the recovery window, and ensured the minimum total relevant costs. They developed an efficient heuristic to solve the problem. In 2015, they developed a simulation model of a three echelon supply chain system with multiple suppliers subject to supply and transportation disruptions [4]. The objective of the paper is to examine the effects of disruption on the system's total recovery costs. Hernn et al. [1] provided a novel simulation-based multi-objective model for supply chains with transportation disruptions, aiming to minimize the stochastic transportation time and the deterministic freight rate.

In this paper, we apply a novel multi-objective programming model for supply chain with transportation disruption. We reference about the method of calculating profit of Lin's paper [5] to study the problem of the manufacturer and retailer's profit when there is a transportation disruption. And we introduce the insurance contract to the model.

3 System Description and Modeling

In the following subsections, we address the system's description of the model. Then present the mathematical representation of the model.

3.1 System Description

In this study, we consider a single supply chain model which has a manufacturer and a retailer. We assume that the information between them is symmetrical. The manufacturer has production, while the retailer has inventory. Before the selling season, the manufacturer and the retailer agree on an insurance contract, but the retailer needs to pay part fees to the manufacturer. And the premium is decided by the insurance company. In our model, we assume that the transportation disruption occurs in the delivery from the manufacturer to the retailer, which interrupts the timely delivery of goods to the retailer. The transportation disruption may be caused by an accident or a natural disaster, such as a earthquake, or flood. In addition, the goods in transit may or may not be damaged during the disruption. Because of the transportation disruption, the insurance contract comes into action.

In this paper, we use a multi-objective programming model to solve the problem. The notations are explained as follows:

Denotes

m: represent the manufacturer;
r: represent the retailer;
I: represent the insurer;
D: the market demand;
π_m: the expected profit of the manufacturer;
π_r: the expected profit of the retailer.

Parameters

x: a random variable presenting the market demand D;
c: a unit cost of the product provided by the manufacturer;
w: a wholesale price that the manufacturer charges the retailer;
p: a retail price that the retailer serves the customers;
h: the shortage cost per unit;
c_t: the transport cost per unit;
s: the salvage value of any unsold product per unit;
α: the retailer's share of losses generated by the deviation of his order quantity from the market demand, and $\alpha \in [0, 1]$, so the manufacturer's share is $1 - \alpha$;
P: the pure premium that the insurer guarantees to pay the manufacturer with the transportation interruption insurance, and it's set by the insurer;
R: the rate of the pure premium to pay to the insurer;
β: the ratio that the retailer takes from the compensation, and $\beta \in [0, 1]$, so the manufacturer takes $(1 - \beta)$L;
$f(x)$: probability density function with the market demand D;
$F(x)$: cumulative density function with the market demand D.

Decision variable

q: an order made by the retailer and the manufacturer based on their forecast of the market demand D.

3.2 Model Formulation

When the transportation disruption occurs, the manufacturer and the retailer's expected profit will be decreased. But with the insurance contract, this situation can be mitigated. So the objective of the insurance contract is to maximize their profit. The objective functions are given as follows:

$$\begin{aligned} \max \pi_m =&(w-c)q - qc_t - (1-\alpha)\left[\int_0^q (p-s)(q-x)f(x)dx \right. \\ &\left. + \int_q^\infty h(x-q)f(x)dx\right] - PR + (1-\beta)P. \end{aligned} \tag{1}$$

Objective function (1) maximizes the manufacturer's expected profit, where $(w-c)q$ is the earnings of the product for the manufacturer with the order quantity from the retailer, qc_t is the cost of transportation, $\int_0^q (p-s)(q-x)f(x)dx + \int_q^\infty h(x-q)f(x)dx$ is the expected losses generated by the deviation of the retailer's order quantity from the market demand, and PR is the cost of the manufacturer paying to the insurer.

$$\begin{aligned} \max \pi_r =&(p-w)q - \alpha\left[\int_0^q (p-s)(q-x)f(x)dx + \int_q^\infty h(x-q)f(x)dx\right] \\ &+ \beta P. \end{aligned} \tag{2}$$

Objective function (2) maximizes the retailer's expected profit, where $(p-w)q$ is the earnings of the product for the retailer while the assumption is that the whole product can be sold out. When the transportation disruption happens, the retailer can get a compensation βP. When the transportation disruption happens, the insurance contract belongs to the third party liability insurance, and its rate is usually a fixed value.

In order to guarantee the integrity of the model, there are constraints as follows:

$$0 < s < c < w < p. \tag{3}$$

If the product have not been sold off, the salvage value of the unsold product s is less than its cost. This is equivalent to a kind of punishment for the exceed quantity. And to make sure for the manufacturer and the retailer's profit, the relationship of unit cost c, the wholesale price w and the retail price p need to be $c < w < p$.

$$0 < h < w. \tag{4}$$

The shortage cost h is less than the wholesale price w.

$$P \geq qc. \tag{5}$$

The purpose to purchase the insurance contract is to minimise the manufacturer and the retailer's profit, so the pure premium P is not less than the cost of the product.

$$alpha \in [0,1], \tag{6}$$
$$beta \in [0,1], \tag{7}$$

where α is the proportion of the expected losses, and β is the proportion of the pure premium.

3.3 Global Model

From the formulation above, a multi-objective model under transportation disruption for the supply chain has been deducted. The aims are to maximize the profit of the manufacturer and the retailer. The global model is given:

$$
\begin{aligned}
\max \pi_m =& (w-c)q - qc_t - (1-\alpha)\left[\int_0^q (p-s)(q-x)f(x)dx \right. \\
& \left. + \int_q^\infty h(x-q)f(x)dx\right] - PR + (1-\beta)P \\
\max \pi_r =& (p-w)q - \alpha\left[\int_0^q (p-s)(q-x)f(x)dx \right. \\
& \left. + \int_q^\infty h(x-q)f(x)dx\right] + \beta P \\
& s.t. \begin{cases} 0 < s < c < w < p \\ 0 < h < w \\ P \geq qc \\ \alpha \in [0,1] \\ \beta \in [0,1]. \end{cases}
\end{aligned} \tag{8}
$$

Theorem 1. *The optimal order quantity for the supply chain system will achieve the maximum only if the following condition is satisfied:*

$$
\alpha^* = \frac{p-w}{p-c-c_t}. \tag{9}
$$

Proof. With the insurance contract, the system's expected profit is

$$
\begin{aligned}
\max \pi = \max \pi_m + \max \pi_r = (p-c-c_t)q - \left[\int_0^q (p-s)(q-x)f(x)dx \right. \\
\left. + \int_q^\infty h(x-q)f(x)dx\right] - PR + P,
\end{aligned} \tag{10}
$$

$$
\frac{\partial \pi}{\partial q} = p - c - c_t - [(p-s+h)F(q) - h] = 0. \tag{11}
$$

Through the function of $\max \pi_m$ and $\max \pi_r$, we can acquire the optimal order quantity q_m^* and q_r^* solving the following equation:

$$
\frac{\partial \pi_m}{\partial q} = w - c - c_t - (1-\alpha)[(p-s+h)F(q) - h] = 0, \tag{12}
$$

$$
\frac{\partial \pi_r}{\partial q} = p - w - \alpha[(p-s+h)F(q) - h] = 0. \tag{13}
$$

Then the optimal order quantity is

$$
q^* = F^{-1}\left[\frac{p-c-c_t+h}{p-s+h}\right]. \tag{14}
$$

The the optimal order quantity of the manufacturer and the retailer is showed by:

$$q_m^* = F^{-1}\left[\frac{(1-\alpha)h + w - c - c_t}{(1-\alpha)(p - s + h)}\right], \tag{15}$$

$$q_r^* = F^{-1}\left[\frac{\alpha h + p - w}{\alpha(p - s + h)}\right]. \tag{16}$$

Generate Eq. (9) into Eqs. (15) and (16), then we can get $q_m^*{=}q_r^*{=}q^*$. So only when $\alpha^* = \frac{p-w}{p-c-c_t}$ is established, the order quantity is to the most optimal value. □

4 Numerical Analysis

In this section, we examine the model by numerical analysis. The data was obtained from the Lin's paper [5], which assumed that the market demand followed uniform distribution $D \sim [400, 500]$. The other parameters were as follows: $p = 18$, $w = 15$, $c = 12$, $s = 8$, $h = 3$, $\beta = 0.6$, $R = 15\%$, and $c_t = 1$.

Take the data into the formulations, and draw the trend chart Fig. 1, which shows that the manufacturer's optimal order quantity increases as α increases, while the retailer's optimal order quantity decreases as α increases.

Let $q_m^*{=}q_r^*$, then we get $\alpha^* = \frac{p-w}{p-c-c_t} = 0.6$. We can find $\alpha^* = 0.6$ from Fig. 1 that proves the model. And from Fig. 1, we can see that the crossing is the optimal order quantity q^* as Table 2, which is equal to 461. This is the best quantity for them. So the pure premium is not less than 5532, and the purchase to the insurer needs to be 830.

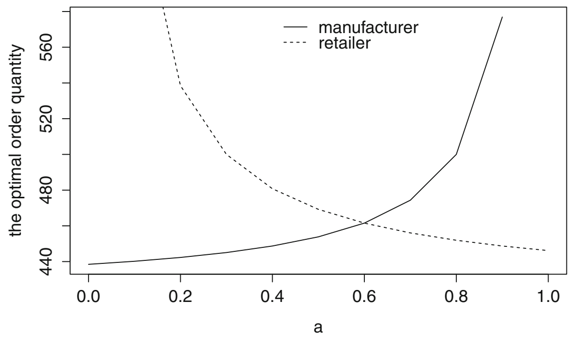


Fig. 1. The effects of α on the optimal order quantity

Table 2. The quantity under different α

Decision	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
m	438.46	440.17	442.31	445.05	448.72	453.85	461.54	474.36	500	576.92	∞
r	∞	653.85	538.46	500	480.77	469.23	461.54	456.04	451.92	448.72	446.15

5 Conclusion

In this paper, we describe the transportation disruption in supply chains. We set a mathematical model which we introduce the insurance contract to deal with transportation disruption the way from the manufacturer to the retailer, and use a case to investigate the applicability of this model. Although the insurance contract is effective in coordinating the supply chain, it also has some limitations. The most critical limitation is that the supplier incurs an administrative cost in monitoring the retailer's sales situation. The objective of the study is to determine the optimal order quantity in the case of the uncertainty of the market demand. This insurance contract transfers the risk from the manufacturer and the retailer to the insurance company, which protects the manufacturer and the retailer's profits and improves the efficiency of the supply chain. In particular, how much to purchase the insurance is discussed in this paper. The model is useful for decision makers to determine the product quantity.

There are several directions for this study to continue. We can extend the model to a complex supply chain with multiple manufacturers or retailers. In addition, we can apply different strategies to deal with transportation disruption.

Acknowledgments. This research was supported by National Natural Science Foundation of China (Grant No. 71640013, and Grant No. 71401020), and System Science and Enterprise Development Research Center (Grant No. Xq16C10), National philosophy and Social Science Foundation of Tianjin (TJGL15-007); Provincial Science and Technology Foundation of Hebei Province (16457643D).

References

1. Chávez H, Castillo-Villar KK et al (2016) Simulation-based multi-objective model for supply chains with disruptions in transportation. Robot Comput Integr Manufact 43:39–49
2. Heckmann I, Comes T, Nickel S (2015) A critical review on supply chain risk—definition, measure and modeling. Omega 52:119–132
3. Hishamuddin H, Sarker RA, Essam D (2013) A recovery model for a two-echelon serial supply chain with consideration of transportation disruption. Comput Ind Eng 64(2):552–561
4. Hishamuddin H, Sarker R, Essam D (2015) A simulation model of a three echelon supply chain system with multiple suppliers subject to supply and transportation disruptions. IFAC Papersonline 48(3):2036–2040
5. Lin Z, Cai C, Xu B (2010) Supply chain coordination with insurance contract. Eur J Oper Res 205(2):339–345
6. Lynch GS (2012) Supply chain risk management 192:33–40. London: Springer
7. Nooraie SV, Parast MM (2015) Mitigating supply chain disruptions through the assessment of trade-offs among risks, costs and investments in capabilities. Int J Prod Econ 171:8–21
8. Shu T, Gao X et al (2016) Weighing efficiency-robustness in supply chain disruption by multi-objective firefly algorithm. Sustainability 8(3):250
9. Zhen X, Li Y et al (2016) Transportation disruption risk management: business interruption insurance and backup transportation. Transp Res Part E 90:51–68

An Interval-Parameter Based Two-Stage Stochastic Programming for Regional Electric Power Allocation

Jingqi Dai and Xiaoping Li(✉)

Business School, Sichuan University,
Chengdu 610065, People's Republic of China
lixiaoping@scu.edu.cn

Abstract. Electricity has become an indispensable part of human being's life. In China, regional electric power allocation, involving multiple groups who pursue the biggest social benefit, has been facing some serious problems and raising people's more and more concern. In this paper, an interval-parameter based two-stage stochastic model is built to solve the problem of planning of electric power allocation target in the perspective of the State Grid Corporation. In the model, the interval-parameter is used to deal with uncertainty of the problem and the two-stage decision making helps take corrective actions after a random events occurs. The application of the model was demonstrated efficiency and practicality with the case study in real world. Combining models proposed in this paper with some related policy implications, it is useful in improving actual operation and helping better decision.

Keywords: Interval-parameter · Two-stage stochastic programming · Electric power allocation

1 Introduction

The second industrial revolution marked the people entered the age of electricity. In recent years, the world electricity industry has experienced a great development and it played a vital role in human beings' life. According to Key World Energy Statistics by International Energy Agency (IEA), world electricity generation increased form 6131 TWh in 1973 to 23816 TWh in 2014, where the growth rate is as high as 288% [8]. However, electrical power systems are facing different problems such as economical and environmental issues. As for the electric power distribution, it is also a complex giant system which has been confusing electricity operators, and researchers have been committed to study it for a long time.

In China, during the last decades, a lasting market-oriented reform has been carried out in electric power industry. Improving the operation efficiency of electric power industry is the primal goal of the government of China, and 2016

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_92

marks the beginning of the second round of such reforms and the 13th Five-year Plan of China [3,4]. Besides, "Further strengthening the institutional reform of the electric power industry was promulgated by the State Council of China whose No. 4 attachment file is about opening electricity plan orderly to promote optimum allocation of electric power resources" [3]. Therefore, on the perspective of the national grid, how to guarantee the power supply and obtain more benefit at the same time are important questions they should face.

In the past years, many scholars studied problem of electric power allocation plan from different perspectives. Many innovative optimization techniques have been developed for allocating and managing electric power in more efficient benign ways [5,7,9,10,13]. Among these methods, linear programming is the basic and prior way to solve this problem. Recently, F. Chen et al [2] used fuzzy chance-constrained programming model to study electric power generation systems planning problem which can reflect uncertain interactions among random variables directly. For power generation systems planning, this approach can have a wider application scope than existing optimization models. However, Huang and his group proposed a new method that is interval-parameter two-stage stochastic program to solve resource management problem, such as water resource management and solid waste management [11,12,14]. This inspired us to apply this kind method to other resource management problem, for example, electric power allocation.

Birge and Louveaux introduced various kinds of two-stage and multi-stage stochastic programming problems [1]. Among them, two-stage stochastic programming is effective for analysing medium-to-long-term planning problem where the system data is characterized by uncertainty [12]. Decision makers will make an initial decision based on the future events and this is called first stage decision. Then, after the uncertainty things are resolved, the second stage decision will be made to correct the results of the first stage decision. Besides, introducing interval parameter into two-stage stochastic model can potentially address uncertainties of the problem.

The goal of the paper is to apply interval-parameter two-stage stochastic program to study problem of electric power allocation while simultaneously taking system uncertainties into account. It will be demonstrated that this kind of method can help decision makers to design more efficient plans and obtain more benefits. The paper is structured as follows: The next section is the statement of key problem. Section 3 presents process of modelling through using the method introduce in this paper. A simple case study was used to demonstrated the effective of the method in Sect. 4. In the end, the paper summarizes the main contents and contributions of this paper in Sect. 5.

2 Key Problem

This paper builds a interval-parameter two-stage stochastic programming model to better solve the problem of electric power allocation. Therefore, it is necessary to introduce some basic background of this problem.

In China, generally speaking, the State Grid Corporation and the government have joint ownership of electric power allocation right. They will make allocation decision overall considering the historical data, requirements of the local government and production capacity of electric power. The authority has an obligation to make a reasonable and scientific electric power allocation plan to achieve the best economic benefits with better resources saving because it will take a lot of resources in the process of electric power production.

Electric power allocation involves a region's residents life, industrial production and other unpredictable situations, so it is a complex system. All the groups need to know how much electric power they can expect because if the guaranteed amount cannot be provided, they have to make a change of their initial planning which may bring them a large loss. For example, the factories have to halt production and the residents have to stop their work that needs use electric power. At the same time, it is also hard for decision makers to make the allocation plan. In practice, there are many stochastic factors that will influence power supply which directly generates to the complexity of the problem. In this situation, this paper uses two-stage stochastic method to solve the problem where the authority will make two decisions. The first decision is to make allocation plan to guarantee basic demands according to historical data and the second decision is to make an compensate that is resulted by the first decision. At the same time, when the State Grid Corporation makes decisions, there are some uncertain factors such as weather, breakdown of machinery that would lead to a gap between demand and supply. Thus, treating uncertain number as interval parameter is a reasonable method to deal with problem's uncertainty.

Therefore, from this discussion, the problem can be solved by building a interval-parameter two-stage stochastic model to obtain more benefits. On one hand, the first decision of insufficient power supply can be made up by the second decision. On the other hand, by adjusting two-phase decision-making, the State Grid Corporation can acquire maximum benefits, which ensures the harmonious development of power industry to a certain extent. The mathematical form for this problem is given in the next section.

3 Modelling

In this section, a interval-parameter two-stage stochastic model with uncertain variables for electric power distribution is constructed. The mathematical description of the problem and the transformation process is given as follows:

3.1 Assumptions

(1) There is no loss during the electric power transmission process;
(2) Interval parameter of the upper and lower values can be obtained through the historical data.

3.2 Notations

To facilitate the problem description, the notations are introduced firstly.
Indices

i: index of provincial power user, where $i = 1, 2, \cdots, n$.

Decision variables

X_i: the first-stage decision variable, which is allocation target for electric power that is promised to user i;
Y_{iS}: the second-stage decision variable, which is shortage of electric power to user i when the actual allocation amount is S.

Uncertain parameters

B_i: net benefit to user i per billion kilowatt hour;
C_i: punish coefficient, which is loss to user i per billion kilowatt hour;
$D_{i\,\min}$: minimal demand of user i to ensure the normal operation of the society;
S_{ih}: the amount where the electric power allocation target is not meet for user i when actual amount of electric power supply is S_i with probabilities p_h,
S: random variable, actual total amount of electric power generation; letting S take values S_h^i with probabilities P_h.

Certain parameters

$S_{i\,\max}$: maximum allowable allocation amount for user i;
p_h: the probability of occurrence of shortage level h, where $h = 1, 2, \cdots, H$, $\sum_{h=1}^{H} p_h = 1$;
$h = 1$: actual power supply is very close to the demand and the shortage is smallest;
$h = 2$: actual power supply is relatively close to the demand and the shortage is medium;
$h = H$: actual power supply is far from to the demand and the shortage is highest.

Symbols

$+$: represent the upper limit of parameters;
$-$: represent the upper limit of parameters.

3.3 The Modeling Process

(1) Objective function

In this paper, we consider regional power grid company has the authority to allocating electric power to multiple users, with the objective of maximizing the total net benefit by optimizing allocation scheme. As these users need to know how much electric power they would be distributed so that they can make sound plans for their production activities to obtain the best social benefit. At the same time, a prescribed amount of electric power is promised to each user according to related provisions because both government and grig company have right to ensure basic function of the society. If the promised demand is met, there will bring benefits to the local economy, otherwise, power shortages could lead to many production activities cannot be carried out, resulting in economic loss. For objective target X_i, the grid company can obtain benefit $\sum_{i=1}^{n} B_i X_i$. However, when uncertain of uncovered electric power appears, the second-stage decision will be made to compensate for any adverse effect which was caused by the first stage of the decision. Therefore, the loss can be expressed as $E\left[\sum_{i=1}^{n} C_i Y_{iS}\right]$, where $E[\cdot]$ represent expected value of a random variable. So, the total net benefits is

$$\max f = \sum_{i=1}^{n} B_i X_i - E[\sum_{i=1}^{n} C_i Y_{iS}].$$

(2) Limitations

Owing to resource constraints and other reasons, there are 4 kinds of constraints and they are detailedly introduced as follows:

① Minimum power demand constraints

Local government has relative policies to promise minimum power demand of the area according to the historical data of electricity consumption so as to ensure normal social life without causing unrest. Based on the above description, we can get the constraint as follow:

$$X_i \geq D_{i\,\min}.$$

② Allowance electric power allocation constraints

For one region, the ability of electric power production is limited. Therefore, allowance allocation amount is also limited so that grid company can't distribute more than actual electric power generation volume to all users. This kind of constraints can be expressed as following:

$$Y_{iS} \leq X_i \leq S_{i\,\max}.$$

③ Available power constraints

In first-stage decision making process, X_i must be determined before actual total electric power supply S are known, while the shortage $Y_i S$ are determined during the second stage when electric power supplies are known but allocation amounts have been fixed. This kind of constraints mean that first-stage decision variable subtract second-stage decision variable should not be more than actual total supply amount. We can have this relationship expressed as:

$$\sum_{i=1}^{n} (X_i - Y_{iS}) \leq S_h^i.$$

④ Non-negative constraints

$$Y_{iS} \geq 0.$$

(3) Overall model

Based on the previous discussion, a two-stage stochastic programming can be built. In this problem, first stage decision must be made before unknown actual total supply amount. When the uncertainty of supply is uncovered, a second-stage decision can be made to compensate influence that is the result of the first-stage decision. We can formulate overall model using two-stage programming:

$$Maxf = \sum_{i=1}^{n} B_i X_i - E\left[\sum_{i=1}^{n} C_i Y_{iS}\right]$$
$$s.t.\begin{cases} X_i \geq D_{i\,\min} \\ Y_{iS} \leq X_i \leq S_{i\,\max} \\ \sum_{i=1}^{n} (X_i - Y_{iS}) \leq S_h^i \\ Y_{iS} \geq 0 \\ \forall i = 1, 2, \cdots, n. \end{cases}$$

(4) Model transformation

There are many uncertain factors in the power system. In real world problem, it is difficult to generate probability distributions for these parameters with small sample size. In this complex giant system, decision makers cannot make accurate prediction of available allocation of electric power in the system. Therefore, the target allocation X_i, net benefit B_i and punish coefficient C_i are uncertain. We introduce the concept of interval parameters to solve the problem of uncertainty, where interval number is defined as a range known upper and lower bounds [6]. Using "+" and "–" represents upper limit and lower limit of parameters respectively. Besides, to solve the problem through liner programming, the allocation S must be approximated as discrete values. In this paper, we let S take value S_i with probabilities $p_h (h = 1, 2, \cdots, H)$. Therefore, we have $E\left[\sum_{i=1}^{n} C_i Y_{iS}\right] = \sum_{i=1}^{n} C_i \cdot E\left[\sum_{i=1}^{n} Y_{iS}\right] = \sum_{i=1}^{n} C_i (\sum_{h=1}^{H} p_h S_{ih})$. Thus, the previous overall model can be converted to the following model:

$$Maxf^{\pm} = \sum_{i=1}^{n} B_{\mathrm{i}}^{\pm} \mathrm{X}_i^{\pm} - \sum_{i=1}^{n} \sum_{h=1}^{H} p_h C_i^{\pm} S_{ih}^{\pm}$$
$$s.t.\begin{cases} \mathrm{X}_i^{\pm} \geq D_{i\,\min}^{\pm} \\ S_{ih}^{\pm} \leq \mathrm{X}_i^{\pm} \leq S_{i\,\max} \\ \sum_{i=1}^{n} (\mathrm{X}_{\mathrm{i}}^{\pm} - \mathrm{S}_{\mathrm{ih}}^{\pm}) \leq S_h^{i\pm} \\ S_{ih}^{\pm} \geq 0 \\ \forall i = 1, 2, \cdots, n. \end{cases}$$

According to research of Wang and Huang [15], this model can be solved through convert interval-parameter two-stage stochastic model into two sub-models which correspond to the upper and lower bounds of the object-function value. It is difficult for decision makers to determine whether the upper bound X^+ or the lower bound X^- of the uncertain variable $X^{\pm}$ correspond to the upper bound of the total net benefit. Thus, maximized total benefit can be obtained by optimized target value. Letting $X_i^{\pm} = X_i^{-} + \Delta X_i z_i$ where $\Delta X_i = X_i^{+} - X_i^{-}$ is a fixed value and z_i is range from 0 to 1. Here, we introduce a new decision variable z_i to identify the optimized target value. Based on above statement, we can get a new transformation model as follows:

$$\max f^{\pm} = \sum_{i=1}^{n} B_{\mathrm{i}}^{\pm}(X_i^{-} + \Delta X_i z_i) - \sum_{i=1}^{n}\sum_{h=1}^{H} p_h C_i^{\pm} S_{ih}^{\pm}$$

$$s.t. \begin{cases} X_i^{-} + \Delta X_i z_i \geq D_{i\min}^{\pm} \\ S_{ih}^{\pm} \leq X_i^{-} + \Delta X_i z_i \leq S_{i\max} \\ \sum\limits_{i=1}^{n} (\mathrm{X_i^{-}} + \Delta \mathrm{X_i z_i} - \mathrm{S_{ih}^{\pm}}) \leq S_h^{i\pm} \\ S_{ih}^{\pm} \geq 0 \\ 0 \leq z_i \leq 1 \\ \forall i = 1, 2, \cdots, n. \end{cases}$$

It can be solved by reformulating above model as two submodels, and the lower bound of objective function value can be first introduced as follows:

$$\max f^{-} = \sum_{i=1}^{n} B_{\mathrm{i}}^{-}(X_i^{-} + \Delta X_i z_i) - \sum_{i=1}^{n}\sum_{h=1}^{H} p_h C_i^{+} S_{ih}^{+}$$

$$s.t. \begin{cases} X_i^{-} + \Delta X_i z_i \geq D_{i\min}^{+} \\ S_{ih}^{+} \leq X_i^{-} + \Delta X_i z_i \leq S_{i\max} \\ \sum\limits_{i=1}^{n} (\mathrm{X_i^{-}} + \Delta \mathrm{X_i z_i} - \mathrm{S_{ih}^{+}}) \leq S_h^{i-} \\ S_{ih}^{+} \geq 0 \\ 0 \leq z_i \leq 1 \\ \forall i = 1, 2, \cdots, n. \end{cases}$$

In addition, the upper bound of objective function value can be expressed as:

$$Max f^{+} = \sum_{i=1}^{n} B_{\mathrm{i}}^{+}(X_i^{-} + \Delta X_i z_i) - \sum_{i=1}^{n}\sum_{h=1}^{H} p_h C_i^{-} S_{ih}^{-}$$

$$s.t. \begin{cases} X_i^{-} + \Delta X_i z_i \geq D_{i\min}^{-} \\ S_{ih}^{-} \leq X_i^{-} + \Delta X_i z_i \leq S_{i\max} \\ \sum\limits_{i=1}^{n} (\mathrm{X_i^{-}} + \Delta \mathrm{X_i z_i} - \mathrm{S_{ih}^{-}}) \leq S_h^{i+} \\ S_{ih}^{-} \geq 0 \\ 0 \leq z_i \leq 1 \\ \forall i = 1, 2, \cdots, n. \end{cases}$$

In the above two models, we can get the optimal solution $S_{ih\mathrm{opt}}^{\pm}$ and $z_{i\mathrm{opt}}$, respectively. Then optimal objective function value is $f_{\mathrm{opt}}^{\pm}$ and it is easy to obtain $X_{i\mathrm{opt}}^{\pm} = X_i^{-} + \Delta X_i z_{i\mathrm{opt}}$. At the same time, optimal electric power allocation is $W_{i\mathrm{opt}}^{\pm} = X_{i\mathrm{opt}}^{\pm} - S_{ih\mathrm{opt}}^{\pm}$.

4 Case Study

A real case study is given in this section to demonstrate the efficiency of this method.

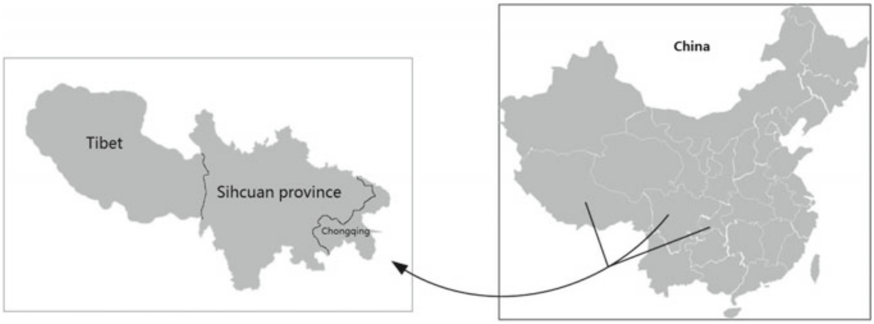


Fig. 1. The location of the case

(1) Description of case

In China, the State Grid Corporation has right to allocate and schedule electric power, which plays an important role in China's power industry. Owning to China covers a large territory, the State Grid Corporate is divided into six sectors in order to better and easier to management, which are North China branch, East China branch, Central China branch, Northeast China branch, Northwest China branch and Southwest China branch, respectively. Each branch has ability to finish their work to provide stable and reasonable power supply and in this process they can obtain benefit. Among them, Southwest China branch is mainly responsible for electric power business of Sichuan province, Tibet and Chongqing city, to manage such work as regional power grid scheduling, operation control and electric power trade.

Southwest China branch is a new sector which was founded in December, 2014. This paper chose Southwest China branch as case to study because this sectors is key to promote economic and social development of southwest of China, and to ensure the safety of national energy efficient supply in strategic height. Besides, this branch only has jurisdiction over the three region which can reduce the computation burden. And the location can be shown in Fig. 1. Therefore, through the case, we want to prove the method used in this paper is feasible.

(2) Data collection and dispose

The data was obtained from home page of State Grid Corporation, government report and market surveys. Table 1 shows the related electric power and economic data which contains demand of different user, punish coefficient and net benefit and so on. Besides, probabilities of degree of electric power shortage and the maximum allowable allocation amount are given in Table 2.

(3) Result analysis

Table 3 shows the results that indicates the most optimal electric power allocation under the different situation of power shortage. It can be seen that these interval solutions stem from uncertainty in input parameters. If the available electric power allocation can't meet the promised amount, there would bring huge loss to the users, resulting in competition among them for the limited

Table 1. Allowable electric power allocation and related economic data

Parameters	Sichuan provinve	Chongqing	Tibet
Net benefit ($B_i^{\pm}$/yuan·kwh^{-1})	[0.3612,0.6015]	[0.510,0.743]	[0.3018,0.6631]
Punish coefficient ($C_i^{\pm}$/yuan·kwh^{-1})	[0.52,0.71]	[0.55,0.76]	[0.39,0.69]
Minimal demand of user ($D_{i\,\min}^{\pm}$/billion kwh)	[2514.79,2697.41]	[606.4,645.87]	[42.23,43.19]
Maximum allowable allocation amount ($S_{i\,\max}$/billion kwh)	3129.44	679.81	44.77
Electric power allocation target ($X_i^{\pm}$/billion kwh)	[2679.64,2991.56]	[631.85,667.93]	[41.25,44.08]

Table 2. The relevant probabilities and actual electric power generation

Probability (p_h)		Low ($h = 1$)	Medium ($h = 2$)	High ($h = 3$)
		0.85	0.10	0.05
Actual total amount of	$S_h^{1\pm}$	[2699.45,2734.21]	[2517.57,2683.82]	[2397.10,2503.78]
Electric power generation	$S_h^{2\pm}$	[627.56,661.08]	[608.49,625.24]	[588.67,605.23]
($S_h^{i\pm}$)/MKh	$S_h^{3\pm}$	[42.13,43.42]	[40.16,42.09]	[38.56,39.97]

electric power. Therefore, the decision making about electric power allocation target plays a vital role in the whole process of the allocation problem. The optimized targets for three uses can be obtained by letting $X_{i\mathrm{opt}}^{\pm} = X_i^{-} + \Delta X z_{i\mathrm{opt}}$, and it is 2679.42 billion kwh, 945.92 billion kwh and 43.2 billion kwh for Sichuan province, Chongqing and Tibet, respectively. At the same time, each shortage and the total benefit can be easy to know form the Table reftab:daijingqi3. From the results, the electric power would be first allocated to Sichuan province, next for Chongqing and the last for Tibet when the scarcity occurs because Sichuan province will bring State Grid Corporation the highest benefit.

In real world, suitable policy is important for decision makers to make plan of an area's sustainable electricity use. From the analysis above, solutions can be obtained by letting the different allocation targets under various policies. Different policy orientation will make the aspect of the consideration vary when decision makers think about the problem. Some policy implications will be given to help making better decision as follows. Firstly, to establish a relatively stable long-term trading mechanism where suppliers and demanders can trade on their own in the competition market. Secondly, improving the mechanism of inter-district electric power deals across the province which permits power generation companies, power users and electricity sell bodies can make occasional trade when electricity is in short supply in a day. Thirdly, environmental protection also is the topic that policy makers need to be concerned with. Thus, forming

Table 3. Electric power allocation solutions under different probabilities

User	Probability	$i = 1$ (Sichuan provence)	$i = 2$ (Chongqing)	$i = 3$ (Tibet)
$S^{\pm}_{ih\text{opt}}$	$h = 1$ (0.85)	0	[6.85,18.36]	[0.66,1.07]
	$h = 2$ (0.10)	[179.83,307.74]	[37.43,42.69]	[1.99,3.04]
	$h = 3$ (0.05)	[300.30,487,78]	[57.25,62.70]	[4.11,4.64]
$X^{\pm}_{i\text{opt}}$		2697.42	945.92	43.2
$z_{i\text{opt}}$		[0.057,1]	[0.39,1]	[0.69,1]
$W^{\pm}_{i\text{opt}}$	$h = 1$ (0.85)	[2697.42,2734.82]	[627.56,661.08]	[42.13,43.432]
	$h = 2$ (0.10)	[2517.57,2683.82]	[608.49,625.24]	[40.16,42.09]
	$h = 3$ (0.05)	[2397.10,2503.78]	[588.67,605.23]	[38.56,39.97]
$f^{\pm}_{i\text{opt}}$		[974.31,1799.42]	[317.56,494.69]	[12.40,29.14]

market mechanism to promote renewable energy use to promise priority use of clean energy.

5 Conclusions

This paper studied interval-parameter two-stage stochastic model applying to the problem of electric power allocation. By using the proposed model, we considered not only how to make a more stable consumer plan, but also let decision makers can benefit more form it. In the process of model building, we overcame uncertainties of the problem through introducing interval-parameter into the model, which at the same time got over complexity of solving process. The initial model was transferred into two submodels so as to obtain solutions, avoiding system failure risk of interval solution. Of course, the approach was demonstrated effective by using data form Southwest branch of the State Grid Corporation of China as a case study. Through analysing result of the case study, some policy implications were proposed in the end which may offer help to regulators of electric power allocation to make decision.

Compared with previous studies, this kind of method is innovative to be applied in the electric power allocation which has good practicality value. However, this method was built based on the assumption that the upper and lower boundary is known. Besides, the probabilities of degree of electric power shortage are also directly given. The future research may focus on these two points to make the study more rigorous, which can better deal with the uncertainty. Besides, some suitable policy simulation also can be put forward to help better decision making.

Acknowledgements. We thank the support of the Social Science Planning Annual Project of Sichuan Province (SC16C011), Soft science research plan project of Sichuan province (2017ZR0098), Soft science research project of Chengdu city (2015-RK00-00156-ZF), The Central University Basic Scientific Research Business Expenses Project of Sichuan University (skqy201623).

References

1. Birge JR, Louveaux F (1997) Introduction to Stochastic Programming. Springer, New York
2. Chen F, Huang GH, Fan YR, Chen JP (2017) A copula-based fuzzy chance-constrained programming model and its application to electric power generation systems planning. Appl Energy 187:291–309
3. CCCCP (Central Committee of the Chinese Communist Party) SSC (2016) Further strengthening the institutional reform of the electric power industry. In Chinese
4. Development NN, Reform Commission NNEA (2016) Six subsidiary documents for the new electric reforms were released. In Chinese
5. Fukuta N, Ito T (2012) A preliminary experimental analysis on combinatorial auction-based electric power allocation for manufacturing industries. In: IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, pp 394–398
6. Huang GH (1998) A hybrid inexact-stochastic water management model. Eur J Oper Res 107(1):137–158
7. Iba K, Suzuki H et al (1988) Practical reactive power allocation/operation planning using successive linear programming. IEEE Trans Power Syst 3(2):558–566
8. Iea IEA (2016) Key world energy statistics 2016. Int Energy Agency 1:37
9. Jordehi AR (2016) Allocation of distributed generation units in electric power systems: a review. Renew Sustain Energy Rev 56:893–905
10. Kayal P, Chanda CK (2016) Strategic approach for reinforcement of intermittent renewable energy sources and capacitor bank for sustainable electric power distribution system. Int J Electr Power Energy Syst 83:335–351
11. Li YP (2014) Municipal solid waste management under uncertainty: an interval-fuzzy two-stage stochastic programming approach. J Environ Inform 12(2):96–104
12. Maqsood I, Huang GH, Yeomans JS (2005) An interval-parameter fuzzy two-stage stochastic program for water resources management under uncertainty. Eur J Oper Res 167(1):208–225
13. Shang JC, Zhang LQ (2007) Research and application of technologies in energy-saving emission-reducing and optimal resource allocation of electric power system. Power Syst Technol 31(22):58–63 In Chinese
14. Wang S, Huang GH (2013) An interval-parameter two-stage stochastic fuzzy program with type-2 membership functions: an application to water resources management. Stochast Environ Res Risk Assess 27(6):1493–1506
15. Wang S, Huang GH (2015) A multi-level taguchi-factorial two-stage stochastic programming approach for characterization of parameter uncertainties and their interactions: An application to water resources management. Eur J Oper Res 240(2):572–581

Pricing Strategies of Closed Loop Supply Chain with Uncertain Demand Based on Ecological Cognition

Dongjing Yu(✉) and Chunxiang Guo

Business School, Sichuan University,
Chengdu 610065, People's Republic of China
1452702457@qq.com

Abstract. With the consumer's ecological cognition becoming more and more higher, remanufacturing has became a trend. It is imperative for contemporary enterprises to find out some ways to balance the profit of supply chain and the government. In this paper, we consider the closed loop supply chain that consists of the manufacturer, retailer and consumers. The retailer is responsible for the acquisition of used products from consumers, and sold them together to the manufacturer. Then the manufacturer producing new products and remanufactured products, they were sold together in the market. Based on the ecological cognition and value judgment of consumers, it is easy to derive the demand quantity of remanufactured products and new products. First, we formulate the cooperative and competitive game model to investigate the manufacturer's and retailer's pricing strategy. Next, in the case of different ecological cognition, we studied the behavior of government subsidy. A mathematical model is presented to formulate the problem to get the optimal decision for the government. Finally, we compare the characteristics of the two models, and draw the corresponding conclusions.

Keywords: Ecological cognition · Cooperative and competitive game model · Closed loop supply chain · Government subsidy

1 Introduction

In the competitive business environment, it is imperative for businesses to consider the resource shortage and environmental pollution. In our country, they are seriously restricting the development of society. Without doubt that recycling of solid waste is an effective way to save resources and reduce pollution. In the recent years, the government has advocated a resource-saving and environmentally friendly mode of production, and resulting in large-scale academic researches on remanufacturing. Therefore, Wei et al. [25] analyzed the obstacles and the incentive factors that should be given in the manufacturing process of China.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_93

In the process of reverse logistics, the factors affecting the recycling of used products are numerous. Pankaj et al. [6] have done a research on the possibility of three way recycling options. Llgin et al. [13] analyzed four major categories about remanufactured products. Fleischmann et al. [8] studied quantitative models for reverse logistics, and [5,11] researched on sustainable development of supply chain systems. In addition, Thierry et al. [24] researched strategic issues on recycling and management product. The spare parts research is also very important, so Ruud H et al. [23] considered the uncertain quality of the used products. Chari et al. [2] addressed a problem about the repairable spare parts to find the cost-optimal production strategy. In order to further study the acquisition price and the quantity of remanufactured products, Shaligram [21] studied the impact of the consolidation center and collection centers. Based on the study of Shaligram et al., Gönsch Jochen [10] supposed that the quantity of used products and replacement parts are unrestricted. Sung et al. [14] devoted to developing optimal production planningnings for different production strategies.

The production demand of remanufactured products is extremely uncertain, so Mitsutaka [17] and others predict the demand of the product by time series analysis. Shi et al. [22] developed an option pricing model to evaluate the acquisition price of used products under uncertain demand and return. For the sake of considering the actual situation of the remanufacturing process will be interrupted, B.C. Giri and S. Sharma [9] studied the complex production planningning under the uncertain demand when the product supply is interrupted. Because there is always a variety of products were produced at the same time, so Shi et al. [22] considered the demand of many kinds of products are uncertain. For solving the problem systematically, Cheng [26] studied the influence of the new product design on the pricing strategy. In order to solve the pricing strategy, a model is established by using the Brown model [16]. Chen et al. [4] used the dynamic programming method to obtain dynamic pricing strategy. Also consider the study of competition, such as Mitra [18] researched the competition strategy under the monopoly industry. Adem et al. [20] researched the competition between an original equipment manufacturer and an independent remanufacture. In the closed-loop supply chain, Chen and Chang [3] considered a remanufacturing models under cooperative and competitive.

Due to the emergence of remanufactured products, many consumers have a different treatment of remanufactured products and new products. Ferrera et al. [7] chosen differentiated prices for new products and remanufactured products. Benjamin et al. [12] considered the role of ambiguity tolerance in consumer perception of remanufactured products. [1,19] consider the impact of government subsidies on the remanufacturig activities. The scholar Li [15] divided the demand of products into two categories by the ecological cognition of consumers, and made a research on the policy of the government subsidy. In this paper, based on the above, we discuss the problem of remanufacturing in closed-loop supply chain, based on ecological cognition and government subsid, then formulate the cooperative and competitive game model to investigate the manufacturer's and retailer's pricing strategy.

This paper is organized as follows: model description is reviewed in Sect. 2 and the proposed model is given in Sect. 3. The key parameters are analyzed in Sect. 4 and the comparison of models will be given in Sect. 5. Finally, some conclusions are summarized in Sect. 6.

2 Model Description

In this paper, we consider three stakeholders, they are manufacturer, retailer and consumer. The manufacturer will produce remanufactured products and new products, and then sold them together to the retailer. The retailer sold remanufactured products and new products to consumers, and collected used products from the consumers, then provide them to manufacturers. The reverse logistics is commissioned by the manufacturer, the manufacturer gives the acquisition price to stimulate the retailers to collect the used products. Thus, the closed-loop supply chain model is shown in Fig. 1.

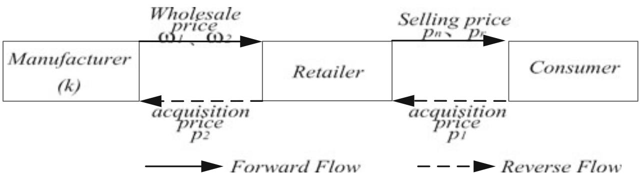


Fig. 1. Model structure of closed loop supply chain

The model structure studied the case of a single manufacturer and a single retailer. Besides, the manufacturer is the leader in the closed loop supply chain. In this paper, we only consider the government takes the corresponding policy mechanism to stimulate the manufacturer. The remanufacturer derives the market demand function of remanufactured products and new products based on the cognitive degree of consumers, then makes the production planning of the product and determines the wholesale price of the product. On the other hand, in order to ensure the orderly conduct of the manufacturing process, the acquisition price is determined by the manufacturers, and the corresponding used product planning is established. Retailers are the followers, according to the manufacturer's wholesale price and acquisition price to determine the retail price of finished product and acquisition price of used products, and optimize their own profit. Finally, we explore the effects of cooperative and competitive model on pricing strategies.

2.1 Notations

The following notations would be used throughout the model

Θ : Ecological cognition of remanufactured products;
p_n : The unit selling price of new product;
p_r : The unit selling price of remanufactured product;
C_1 : The unit cost of sales for new product;
C_2 : The unit cost of sales for remanufactured product;
p_1 : The unit acquisition price of the retailer to user;
p_2 : The unit acquisition price of the manufacturer to the retailer;
C_r : The unit production cost of remanufactured product;
C_n : The unit production cost of new products;
k : Government subsidies for the remanufacturer productions;
q_r : The demand quantity of remanufactured products;
q_n : The demand quantity of new products;
ω_r : The unit wholesale price of remanufactured products;
ω_n : The unit wholesale price of new product;
Π_r : The profit function for retail;
Π_n : The profit function for manufacturer;
Π : The profit function for the supply chain.

2.2 Assumptions

Assumption 1. The decision of the consumer to purchase the product is based on the utility of the product. Then we supposed the consumers' evaluation of new products is v, and supposed the v is uniformly distribution, its probability density function is:

$$f(x) = \begin{cases} 1, \, x \in [0,1] \\ 2, \, x \notin [0,1] \, . \end{cases}$$

Supposed that the market size is 1, and each consumer only buy one product.

Assumption 2. We defined the θ is the degree that consumer based on environmental awareness and willing to buy the remanufactured products. We supposed $0 \leq \theta \leq 1$, when $p_n > \theta v$, customers will not buy new products, when $p_r > \theta v$, customers will not buy remanufactured products. And if $p_n \leq v$, $p_n > \theta v$, meanwhile, $\theta v - p_r \leq \theta v - p_n$, consumers will choose the new products. When $p_n \leq \theta v$, $p_r \leq \theta v$, and $v - p_r \leq \theta v - p_n$, consumers will choose the remanufactured products.

Assumption 3. Due to make the production continue, here, we assume that $p_1 \leq p_2, C_r \leq C_n$.

Assumption 4. Manufacturers and retailers are based on complete information symmetry, it means that each of them fully learn of the cost of each other, pricing, strategy, and other relevant information, so that the next Steinberg game model can work well.

Assumption 5. Suppose that in the process of remanufacturing, the rate of production is 100%. It means that no waste products in the process of manufacturing. And used production can meet the market demand.

Assumption 6. According to the degree of consumers' cognition, the market demand function of remanufactured product and new products is derived:

$$q_n = 1 - \frac{p_n - p_r}{1-\theta}, \quad \frac{p_r}{\theta} \le p_n \le p_r + 1 - \theta,$$
$$q_r = \frac{\theta p_n - p_r}{\theta(1-\theta)}, \quad p_n + \theta - 1 \le p_r \le \theta p_n.$$

3 Model Development

Based on the above, we obtain the profit function of the manufacturer and the retailer, the profit function of manufacturer:

$$\pi_n = (\omega_n - c_n)q_n + (\omega_r - c_r - p_2)q_r + kq_r. \tag{1}$$

The first element on the right side of the equation is the profit of the manufacturer for selling new products; The second element on the right side of the equation is the profit of the manufacturer for selling remanufactured products; the third element on the right side of the equation is the profit of the manufacturer produces the remanufactured products and then gets the government subsidies.

The profit function of retailer:

$$\pi_r = (p_n - \omega_n - c_1)q_n + (p_r - p_1 - c_2 - \omega_r + p_2)q_r. \tag{2}$$

The formula on the right side of the first represent the profits of retailer for selling new products, of which second is stand for the profits for selling remanufactured products. So there is a total profit of the closed-loop supply chain:

$$\pi = \pi_n + \pi_r = (p_n - c_n - c_1)q_n + (p_r - p_1 - c_r - c_2 + k)q_r. \tag{3}$$

In this paper, we consider the following two cases in the closed loop supply chain.

3.1 Competitive Pricing Strategy

In the competitive pricing strategy, the manufacturer and the retailer take their own benefit maximization as the decision-making goal respectively. In this decision models, the manufacturer is in the dominant position, and the retailer is the follower. So the paper uses the reverse induction method to solve the model. The derivation of the above formula, we can get:

$$\begin{bmatrix} \partial^2\pi_r/\partial^2 p_n = 2/(\theta-1) & \partial^2\pi_r/\partial p_n \partial p_r = -2/(\theta-1) \\ \partial^2\pi_r/\partial p_r \partial p_n = -2/(\theta-1) & \partial^2\pi_r/\partial^2 p_r = 2/(\theta(\theta-1)) \end{bmatrix}.$$

In order to verify whether there is the optimal solution of the equation, we need to calculate the characteristic value of Hessian matrix and judge that they are negative. After that we get Matrix eigenvalue:

$$\begin{bmatrix} -(\theta-(5\theta^2-2\theta+1)^{(1/2)}+1)/(-\theta^2+\theta) & 0 \\ 0 & -(\theta+(5\theta^2-2\theta+1)^{(1/2)}+1)/(-\theta^2+\theta) \end{bmatrix}.$$

Next, we just need to prove it: $(\theta-(5\theta^2-2\theta+1)^{(1/2)}+1)/(-\theta^2+\theta)>0$. And then we can get the conclusion that the Hessian matrix is negative, so there is a corresponding maximum value. Due to the $\theta\in[0,1]$, the denominator is greater than zero, it is necessary to prove that the molecule is greater than zero. Easy to get $4\theta-4\theta^2$, so in this model, the retailer has the optimal retail price. We take the partial derivative of Eq. (2):

$$\partial\pi_r/\partial p_n=0\partial\pi_r/\partial p_r=0.$$

And obtain that:

$$p_n=(c_1-c_2-\theta-p_1+p_2+2p_r+\omega_n-\omega_r+1)/2, \tag{4}$$
$$p_r=(c_2+p_1-p_2+\omega_r-c_1\theta+2\theta p_n-\theta\omega_n)/2. \tag{5}$$

Using formula (4) and (5), we can get:

$$p_n=(c_1+\omega_n+1)/2, \tag{6}$$
$$p_r=(c_2+\theta+p_1+\omega_r-p_2)/2. \tag{7}$$

Put p_r and p_n into the Π_n, and then consider the equation:

$$\begin{matrix}\partial\pi_n/\partial\omega_n=0,\\ \partial\pi_n/\partial\omega_r=0\end{matrix}.$$

At last, we can obtain that:

$$\omega_n^*=(1/2)(c_n+1-c_1), \tag{8}$$
$$\omega_r^*=(1/2)(c_r-c_2+\theta-k-p_1+p_2). \tag{9}$$

Combine Eqs. (6), (7), (8) and (9), we can get that:

$$p_n^*=(1/4)(c_1+c_n+3), \tag{10}$$
$$p_r^*=(1/4)(c_2+3\theta-k+p_1-p_2+c_r). \tag{11}$$

In this paper, use * to represent the optimization.

3.2 Cooperative Pricing Model

Under this model, the manufacturer and the retailer take the Π as the objective function. This problem can be translated to the maximum value of the equation:

$$\pi=\pi_n+\pi_r=(p_n-c_n-c_1)q_n+(p_r-p_1-c_r-c_2+k)q_r. \tag{12}$$

The Hessian matrix of the equation is obtained as follows:

$$\begin{bmatrix} \partial^2\pi/\partial^2 p_n = 2/(\theta-1) & \partial^2\pi/\partial p_n \partial p_r = -2/(\theta-1) \\ \partial^2\pi/\partial p_r \partial p_n = -2/(\theta-1) & \partial^2\pi/\partial^2 p_r = 2/(\theta(\theta-1)) \end{bmatrix}.$$

In the same way, we can know that there is a optimal solution. And solved it, we can get that:

$$p_r^* = (1/2)(c_2 + c_r + \theta - k + p_1 + p_2), \tag{13}$$

$$p_n^* = (1/2)(c_1 + c_n + 1). \tag{14}$$

4 Numerical Example

In order to describe the effect of θ on p_n, p_r, q_n, q_r, we fixing the other parameters as the base setting: $C_1 = 0, C_2 = 0.05, C_n = 0.2, C_r = 0.15, k = 0.04, p_1 = 0.05$ and $p_2 = 0.08$. Then we made the Tables 1, 2, and compared the different in the two model. By the way, and in the next graphic analysis, if there is no special note, the value of the parameters as the same as the table. From the following table: In the cooperative pricing model, the government's regulation should be greatly strengthened. Because in the Table 2, the number of remanufactured products appear negative growth, and their sales prices are gradually increasing. It means that the remanufactured products is no profit at all, the only way is to continuously improve the sales price of remanufactured products to get profit. As can be seen from the table, the sales price of new product does not change with the increase of θ. The reason is that we use the reverse induction method to solve this model. The retailers is the first to determine the market price, then manufacturer get the price of wholesale. As a result, the sales price of new products is only related to the cost of sales and production cost of new products.

First, we make some graphic analysis in the model of competition. The effect of θ on market demand and market profits is now studied and the results are shown in Figs. 2 and 3 respectively.

For the analysis of the Fig. 2, we have the following conclusions: at the intersection, when the $\theta = 0.8781$, the market demand for new and remanufactured

Table 1. The effect of θ on parameter in competition

Variable parameter	θ Value of the corresponding parameter variation								
θ	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.0
q_r	−0.0104	0.0000	0.0119	0.0267	0.0469	0.0784	0.1389	0.3158	−Inf
q_n	0.2063	0.2000	0.1917	0.1800	0.1625	0.1333	0.0750	−0.1000	Inf
p_r	0.5200	0.5575	0.5950	0.6325	0.6700	0.7075	0.7450	0.7825	0.8200
p_n	0.8000	0.8000	0.8000	0.8000	0.8000	0.8000	0.8000	0.8000	0.8000

Table 2. The effect of θ on parameter in cooperation

Variable parameter	θ Value of the corresponding parameter variation								
θ	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.0
q_r	−0.3542	−0.3516	−0.3571	−0.3733	−0.4063	−0.4706	−0.6111	1.0526	−Inf
q_n	0.6125	0.6286	0.6500	0.6800	0.7250	0.8000	0.9500	1.4000	−Inf
p_r	0.4450	0.4700	0.4950	0.5200	0.5450	0.5700	0.5950	0.6200	0.6450
p_n	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000

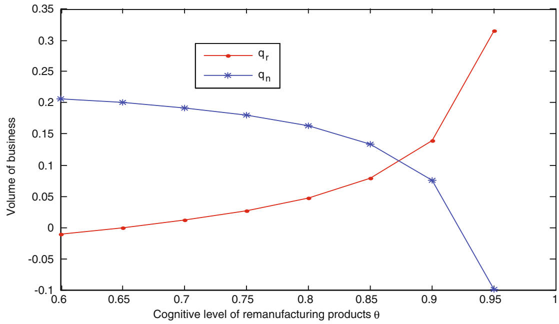


Fig. 2. The effect of θ on market demand.

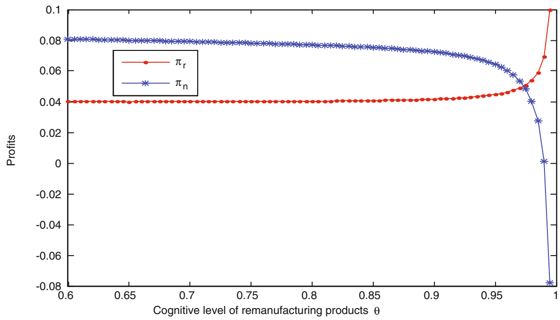


Fig. 3. The effect of θ on market profits

products is equal. At this case, when $\theta > 0.8781$, consumers tend to buy manufactured products. Manufacturer can only producing the remanufactured products when the remanufactured product is profitable enough. When $\theta = 0.9300$, manufacturers no longer produce the new products, because at this point, the demand for new products is zero. Why it happened? Maybe there is no bias towards the new products and remanufactured products, and the price of the remanufactured product is lower than the new product. In Fig. 3 we can draw that: when $\theta = 0.9735$, the remanufacturer and the retailer's profit is equal, and in this mode, the manufacturer produces remanufacturer products are negative profits. As in Fig. 2, with the increase number of remanufacturer products, the manufacturer's profits have been negative growth. And it is found that they can produce remanufacturer products to adjust the profit distribution between manufacturer and retailer. When θ is larger than 0.9901, the manufacturer is no longer profitable, if retailer and government do not adjust it in this case, the mode of production will no longer continue. It can be seen that the remanufacturing does not allow manufacturers to make a profit. As a result, the main incentive for manufacturers to make the profit will be to improve remanufacturing technology, save remanufacturing costs and get more government subsidies.

Here, we consider the situation: when the degree of ecological cognition is very low, in order to advocate environmental protection activities, how should the government make the decisions? We considered the following three cases, see the chart below.

From the Fig. 4, it is observed that when $\theta = 0.6000$, $k_1 = 0.05$, the retailer's profit appears to the minimum value. With the increase of K, which can reduce the number of new product and increase quantity of remanufactured products. The wholesale price of products will be reduced with the increase in government subsidies. As a result, the unit selling price of remanufactured product also begins to reduce. Due to the decrease in the number of new product, retailers' profits

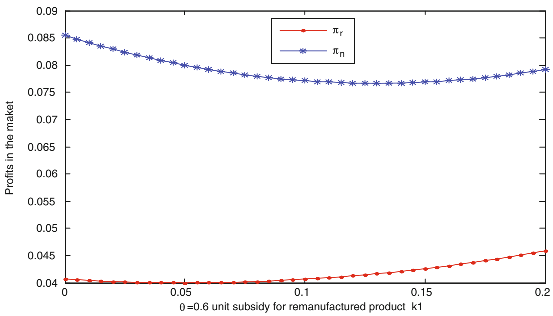


Fig. 4. The impact of government subsidies on market profit when $\theta = 0.6$

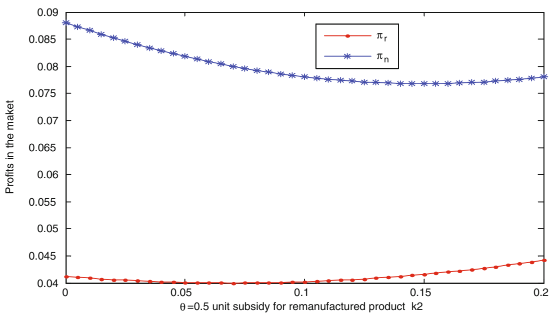


Fig. 5. The impact of government subsidies on market profit when $\theta = 0.5$

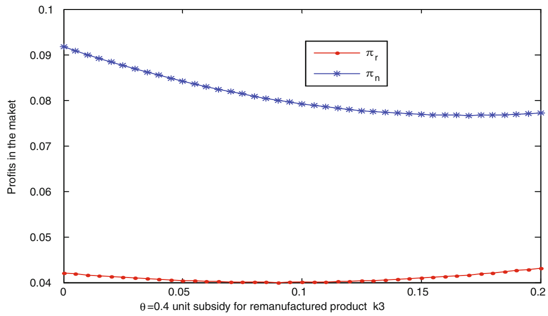


Fig. 6. The impact of government subsidies on market profit when $\theta = 0.4$

continue to decline. Only when the government subsidy is greater than 0.05, the retailer's profit began to grow. We can find the selling profits mainly come from new products. The pr decreases with the k_1 increase, that makes the consumer tends to buy remanufactured products. As a result, the demand of new product begin to decline. When the $k = 0.13$, because the government subsidies make up for the loss in remanufacturing, the manufacturers' profit function begins to grow, so the retailer's selling price also increased significantly. From the Figs. 4, 5 and 6 we can draw that: in the case of other variables are constant, the smaller the value of θ, the more serious decline in the profit function. It should be noted that with the decrease θ the lowest point of red and blue line start move to the right. We can draw that the higher cognition of remanufactured products, the government's subsidy will be smaller. The model also reveals that manufacturers and retailers will not be able to make a remanufacturing activity if the government does not take the policies and give subsidies.

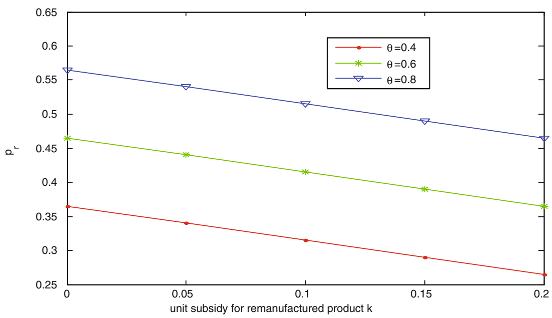


Fig. 7. The variation of p_r with respect to K under different θ

Next, we make some graphic analysis in the model of cooperation. In order to make more profits, the manufacturer and retailer set prices together. We assumed that other factors are constant, and then discuss the k if the government wants

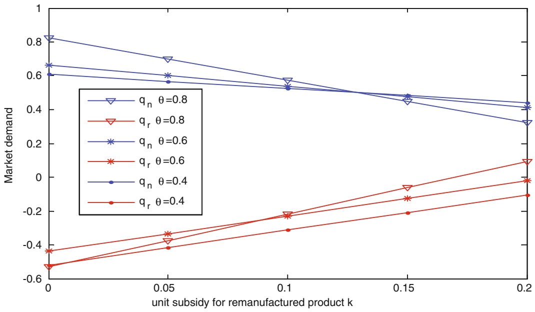


Fig. 8. The variation of market demand with respect to K under different θ

to make remanufacturing continue. From the previous discussion, we learn that the change of k does not affect the p_n. So we compared the effect of k on p_r in the following cases. As follows: from the Fig. 7, we can clearly see a linear relationship between the p_r and k, the ratio is -0.5. The more greater the k, the more smaller the unit retail price of remanufactured products. At the same time, we can also find that in the process of pricing, government subsidies are particularly important. At the same time, we also compared the demand function of new products and remanufactured products in the Fig. 8. It is obvious that when the θ is bigger, the demand market of remanufactured products or new products are more sensitive to the k. Because with the increase of θ, the more steep line. When the $\theta = 0.8$, the government needs to provide at $k = 0.17$, because the quantity of remanufactured products is zero at this point of $k = 0.17$. Only when k0.17, the closed loop supply chain can continue to remanufacturing. For the sake of reducing the subsidy, the government can do the following adjustments: to enhance ecological cognition of remanufactured products and increase the value of θ. Because in the Fig. 8, we can know that the θ is more bigger, the possibility is more higher for manufacturer to remanufacturing. Secondly, our government should establish a complete production system of closed loop supply chain, it can reduce the unnecessary cost waste. Besides, remanufacturing costs account for a large proportion of the cost of production, and it seriously hinders the development of remanufacturing. It is necessary for government to provide technical support for the remanufacturing. At last, if the government can destruct the joint policy of manufacturers and retailers, it will take profit to him. Because in the joint pricing strategy, the new products can bring more profits to them.

5 Comparison of Models

In this paper, we study cooperative and competitive game model in the closed loop supply chain. In the model of competition, due to the manufacturers and retailers make decisions respectively, there is competition between them. Producing the remanufactured products can adjust the interests of the relationship between the manufacturer and the retailer. It is benefit for the government to promote remanufacturing. And the retailer more like to make remanufactured products to compete interests with manufacturer. But From Tables 1 and 2 we can obviously find remanufacturing hurts the interests of manufacturer, retailer and consumer. So them will choose cooperative pricing model. In the Fig. 7, it is not difficult to find that the total profit of cooperation is always higher than the competition. Producing remanufactured products could not make profits, so the profits always decrease with the increase of the degree of cognition. When the θ from 0.9 to 0.95, there is a sharp decline in the total profit of the two models. The reason is the rapidly decrease in the number of new products and the number of remanufactured products rapidly increase. It proves again that produces the remanufactured products will not bring profit to manufacturers. The main reason of remanufacturing could not make profits is that the remanufacturing technology has not kept pace with the society. It also reveals that if the government not take any actions, it will happen nothing on remanufacturing.

6 Conclusion

In this paper, we consider the closed loop supply chain that consists of the manufacturer, retailer and consumers, discuss two models of cooperative and competitive game model. At the same time, we using the game theory to analysis the model and the conclusions are the following.

In the cooperative and competitive game model, if the government no support or no provide subsidy, remanufacturing will not be able to carry on. And in the cooperative game model, the government's regulation should be strengthened. Because the cooperation is not conducive to producing remanufactured products. As the number of new products and profits are declining, manufacturers have a reason to down the price of new products to get more quantity. But the optimal wholesale price in this paper has nothing to do with θ, maybe in this model we use the reverse induction method to solve the model, the retailers is the first to determine the market price, then manufacturer make the price of wholesale. What's more, producing remanufactured products can adjust the profit distribution between manufacturer and retailer. And in cooperation, it can bring higher profit to the supply chain. But it does not accord with the interests of the government. In order to save economic costs, the government is more willing to see them make decisions separately. With the θ raised, it will not good for manufacturers and retailers, but it is good for the government to take actions. The discussion of these two models is very important for our research in the future, and give the direction of the government to advocate green environmental protection. In the future work, we could consider the complex closed loop supply chain systems with multiple manufacturers and multiple retailers, and the used products will be pre-sale to the manufacturers. What's more, used products will be divided into different grades by the consolidation center. The consolidation center will be the new participator in the future model.

Acknowledgements. The authors wish to thank the Major International Joint Research Program of the National Natural Science Foundation of China (Grant No. 71020107027), the National Natural Science Foundation of China (Grant No. 71001075), Doctoral Fund of Ministry of Education of China (No. 20110181110034), and Social sciences program of Chengdu (No. SKNHR13-07), Central University Fund of Sichuan University (No. skqy201112) under which the present work was possible.

References

1. Aksen D, Aras N, Karaarslan AG (2009) Design and analysis of government subsidized collection systems for incentive-dependent returns. Int J Prod Econ 119(2):308–327
2. Chari N, Diallo C et al (2016) Production planning in the presence of remanufactured spare components: an application in the airline industry. Int J Adv Manuf Technol 87(1):957–968
3. Chen JM, Chang CI (2012) The co-opetitive strategy of a closed-loop supply chain with remanufacturing. Transp Res Part E Logist Transp Rev 48(2):387–400
4. Chen JM, Chang CI (2013) Dynamic pricing for new and remanufactured products in a closed-loop supply chain. Int J Prod Econ 146(1):153–160
5. Choi TM, Li Y, Xu L (2013) Channel leadership, performance and coordination in closed loop supply chains. Int J Prod Econ 146(1):371–380
6. Dutta P, Das D et al (2016) Design and planning of a closed-loop supply chain with three way recovery and buy-back offer. J Clean Prod 135:604–619
7. Ferrer G, Swaminathan JM (2010) Managing new and differentiated remanufactured products. Eur J Oper Res 203(2):370–379
8. Fleischmann M (2004) Quantitative models for reverse logistics. Springer, Heidelberg
9. Giri BC, Sharma S (2015) Optimal production policy for a closed-loop hybrid system with uncertain demand and return under supply disruption. J Clean Prod 112:2015–2028
10. Gönsch J (2015) A note on a model to evaluate acquisition price and quantity of used products for remanufacturing. Int J Prod Econ 169:277–284
11. Guide VDR, Wassenhove LNV (2009) The evolution of closed-loop supply chain research. Oper Res 57(1):10–18
12. Hazen BT, Overstreet RE et al (2012) The role of ambiguity tolerance in consumer perception of remanufactured products. Int J Prod Econ 135(2):781–790
13. Ilgin MA, Gupta SM (2010) Environmentally conscious manufacturing and product recovery (ecmpro): a review of the state of the art. J Environ Manage 91(3):563–591
14. Jung KS, Dawande M et al (2016) Supply planning models for a remanufacturer under just-in-time manufacturing environment with reverse logistics. Ann Oper Res 240(2):1–49
15. Li XQ (2015) Research on the policy of remanufacturing subsidy based on the ecological cognition of consumers. Prod Res 2015(9):12–19 in Chinese
16. Liang Y, Pokharel S, Lim GH (2009) Pricing used products for remanufacturing. Eur J Oper Res 193(2):390–395
17. Matsumoto M, Komatsu S (2015) Demand forecasting for production planning in remanufacturing. Int J Adv Manufact Technol 79(1):161–175
18. Mitra S (2015) Models to explore remanufacturing as a competitive strategy under duopoly. Omega 20:215–227

19. Mitra S, Webster S (2008) Competition in remanufacturing and the effects of government subsidies. Int J Prod Econ 111(2):287–298
20. Örsdemir A, Parlaktürk AK (2014) Competitive quality choice and remanufacturing. Soc Sci Electron Publ 23(1):48–64
21. Pokharel S, Liang Y (2012) A model to evaluate acquisition price and quantity of used products for remanufacturing. Int J Prod Econ 138(1):170–176
22. Shi J, Zhang G, Sha J (2011) Optimal production planning for a multi-product closed loop system with uncertain demand and return. Comput Oper Res 38(3):641–650
23. Teunter RH, Flapper SDP (2011) Optimal core acquisition and remanufacturing policies under uncertain core quality fractions. Eur J Oper Res 210(2):241–248
24. Thierry M, Salomon M, Nunen JV, Wassenhove LV (1995) Strategie issues in product recovery management. Calif Manag Rev 37(2):114–135
25. Wei S, Cheng D et al (2015) Motives and barriers of the remanufacturing industry in China. J Clean Prod 94:340–351
26. Wu CH (2012) Product-design and pricing strategies with remanufacturing. Eur J Oper Res 222(2):204–215

Beds Number Prediction Under Centralized Management Mode of Day Surgery

Jianchao Yang[1], Luo Li[1], Hongsheng Ma[2], and Yong Luo[1](✉)

[1] Business School, Sichuan University, Chengdu 610064, People's Republic of China
scuyangjc@163.com
[2] West China Hospital, Sichuan University, Chengdu 610041, People's Republic of China

Abstract. In this paper, we firstly use simulation methods to reach the conclusion that the beds allocation scheme under centralized management mode is more advantageous to either elective surgery patients or day surgery patients, and the beds allocation scheme under decentralized mode is more equitable. Then we do a large number of experiments to explore the beds number allocate to day surgery patients under centralized management mode, and seek the allocation scheme which is advantageous to day surgery patients and the equity is closest to the same situation under decentralized mode. At last, based on the data we obtain from the experiments, regression and neutral network are used to predict the beds number allocated to day surgery patients under centralized management mode.

Keywords: Day surgery · Beds number prediction · Linear regression · Neural network

1 Introduction

There are three types of surgery management modes adopted overseas: day surgery center inside a hospital, free-standing day surgery center and operation room at the clinic [7]. The mainly adopted mode in our country is day surgery center inside a hospital under three different management modes: centralized management mode, decentralized management mode and the management mode combining centralized mode and decentralized mode. Centralized management mode is an integrated management mode where day surgery center, as the centralized management platform, assembles all patients together to conduct centralized admission, scheduling and follow-up visit. While decentralized management mode is a mode where day surgery is managed by departments.

In this paper, we study hospitals under the mode where centralized and decentralized management coexist. Under centralized management mode, day surgery patients are admitted in day surgery center where their admission, operation and discharge are arranged. Day surgery center is self-contained with wards

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_94

and operation room. The beds and operation rooms of day surgery patient are dedicated, not shared with elective patents. Under decentralized mode, wards and operation rooms of day surgery and elective surgery are managed by departments. The beds and operation rooms has no boundary between day surgery patients and elective surgery patients, i.e. to some extent, beds and operations are shared by day surgery patients and elective surgery patients.

We conduct simulation study on centralized and decentralized management modes in this paper. Through a large number of simulation experiments, we reach the conclusion that centralized management mode is more advantageous to either elective surgery patients or day surgery patients. And decentralized management mode is more equitable. While the development of day surgery is encouraged, centralized management mode is a better choice. But if beds allocation is far from equitable under centralized management mode, it will be exceedingly advantageous to one of the two types of patients. Thus, in this paper, we seek the beds number allocate to day surgery patients, under Centralized management mode, that is advantageous to day surgery patients and has Equity Closest to the same situation under Decentralized mode (CECD). Based on the large amount of simulation data, we screen out characteristic variables and use regression and neural network to predict the beds number allocate to day surgery patients under CECD. And then we verify the effectiveness of the regression equation and neural network model.

Technological and theoretical advances in computer science and mathematics offer new options to complement traditional statistical analysis [7]. A major focus of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. In the past decade, machine learning algorithms have revealed previously undetected trends in historical data [9,13]. Especially, artificial neural network and traditional linear regression are effective method. There are some studies use linear regression and neural network to predict the objective function value. Such as, Menke et al. [7] designed an artificial neural network to predict emergency department volume. Shi et al. [11] performed hierarchical linear regression and propensity score matching to test hospital/surgeon volume for associations with breast cancer surgery costs. Tsai [12] developed artificial neural network models to predict length of stay for inpatients with one of the three primary diagnoses: coronary atherosclerosis, heart failure, and acute myocardial infarction in a cardiovascular unit in a Christian hospital in Taipei, Taiwan. Li et al. [15] proposed an artificial neural network model to predict the severity of menopausal symptoms. Launay et al. [5] used artificial neural networks to predicate of prolonged length hospital stay in older patients hospitalized in acute care wards after an emergency department. Gholipour et al. [4] used a neural network for predicting survival and length of stay of patients in the ward and the intensive care unit of trauma patients and to obtain predictive power of the current method. Wise et al. [14] used artificial neural network model to provide vascular surgeons a discriminant adjunct to assess the likelihood of in-hospital mortality on a pending Ruptured abdominal aortic aneurysm admission. Some studies use linear regression and neural network to study the influencing factors of disease or identify influencing factors. Such as, Shakerkhatibi et al. [10] were conducted to evaluate the relationship between air pollutants and hospital admissions for cardiovascular and respiratory diseases using the artificial neural networks and conditional logistic regression modeling. Abdullah [1] aimed to model the relationship between Health Related Quality of Life variables using an integrated model of fuzzy inference system and linear regression. Bonellie [3] used regression and logistic regression models to investigate changes over time in size of babies particularly in relation to social deprivation, age of the mother and smoking. Some studies use linear regression and neural network do clustering and classification, Aguiar et al. [2] developed one artificial neural network model for classification and another risk group assignment for Pulmonary tuberculosis in hospitalized patients. And some studies focus on the methods improvement. Russell et al. [8] through their research discovered that bayesian linear regression provides a significantly more accurate estimate of the rate of change in mean sensitivity than the standard ordinary least squares linear regression approach.

Artificial neural network and traditional linear regression can be well applied to the prediction of the objective function value, the identification of the influencing factors, clustering and classification. In our paper, we use linear regression and neural network to predict the objective function value, the beds number under CECD.

2 Evaluation of Two Management Mode of Day Surgery Based on Simulation

2.1 Parameters

In this paper, taking Laparoscopic Cholecystectomy (LC) as an example, we build two simulation models. Model I is the centralized management model, where beds and operation rooms occupied by day surgery patients and elective surgery patients are independent. And there are dedicated beds and operation rooms for day surgery patients. That is day surgery patients wait in a queue for admission and operation according to the First Come and First Served (FCFS) policy, they can only occupy beds and operation rooms dedicated to day surgery patients on admission. And elective surgery patients wait in other queue for admission and operation according to the FCFS policy, they can only occupy beds and operation rooms dedicated to elective surgery patients on admission. Model II is the decentralized management mode, where beds and operation rooms occupied by day surgery patients and elective surgery patients are shared, i.e. day surgery patients has no dedicated bed or operation room. Day surgery patients and elective surgery patients wait in one queue for admission and operation according to the $FCFS$ policy.

Significant input parameters of the model are as follows.

B : total number of beds for day surgery patients and elective surgery patients;
B_D : number of beds for day surgery patients;
B_E : number of beds for elective surgery patients, $B_D + B_E = B$;
λ_D : arrival rate of day surgery patients (unit: person per day);
λ_E : arrival rate of elective surgery patients (unit: person per day);
OT_D : operation time of day surgery patients (unit: hour);
OT_E : operation time of elective surgery patients (unit: hour);
RT_D : length of hospital stay after operating day surgery (unit: hour);
RT_E : length of hospital stay after operating elective surgery (unit: hour);
WT_1 : average waiting time before admission (unit: hour), in this paper we define
$WT_1 =$ time of admission $-$ time of arrival;
A_D : number of day surgery patients arrived (unit: person);
A_E : number of elective surgery patients arrived (unit: person);
S_D : number of day surgery patients served (unit: person);
S_E : number of elective surgery patients served (unit: person);
S : total number of surgery patients served (unit: person), $S = S_D + S_E$

2.2 Simulation Results and Analysis

In this paper, we define the regular opening hour of each operation room as 8 h, at the end of each workday, operations still in progress should be finished with extra work, newly-opened operations are not allowed, and operation rooms are closed for weekends. When we conduct simulation in this paper, OT_D, RT_D, OT_E and RT_E are considered as random numbers generated according to the mean values and standard deviations in Table 1.

Table 1. Descriptive statistics result [6]

	Surgery type	N	Mean value	Min	Max	SD	SE mean
OT	Elective surgery	1841	72.2819	20	1116	56.92553	1.32672
	Day surgery	682	64.0176	20	674	33.30224	1.27521
RT	Elective surgery	1841	2696.3781	920	43285	2391.37266	55.73402
	Day surgery	682	1198.2361	600	8355	713.50039	27.32134

The simulation period of Model I and Model II are both 1 year, and the conclusion we reach through a large number of simulation experiments are in accordance. In this paper we only elaborate on the simulation result when $_D = 8$, $_E = 30$. The number of elective surgery patients arrived (A_D), number of elective surgery patients arrived (A_E), number of day surgery patients served (S_D), number of elective surgery patients served (S_E), total number of surgery

Table 2. Simulation result when $\lambda_D = 8$, $\lambda_E = 30$ (simulation period: 1 year, approximately 260 days)

Model	$A_D(p)$	$A_E(p)$	$S_D(p)$	$S_E(p)$	$S(p)$	$WT_1(h)$
Model I ($B_D = 4$)	2080	7800	1118	4324	5442	1397
Model I ($B_D = 5$)	2080	7800	1404	4247	5651	1308
Model II	2080	7800	1148	4306	5454	1382

patients served (S), average waiting time before admission (WT_1) are shown in Table 2.

According to Table 2, under the condition where $B_D = 4$, Model I has 30 fewer day surgery patients served than Model II, 18 more elective surgery patients served than Model II. To sum up, Model I has 12 fewer patients served than Model II, and 15 h longer average waiting time before admission than Model II.

While, under the condition where $B_D = 5$, Model I has 256 more day-surgery patients served than Model II, 59 fewer elective-surgery patients served than Model II. To sum up, Model I has 197 more patients served than Model II, and 74 h shorter average waiting time before admission than Model II.

The result shown in Table 2 indicates that, under centralized management mode, bed and operation room resources of day surgery and elective surgery are respectively independent, and it is advantageous either to elective surgery patients (under the condition where $B_D \leq 4$ in Table 2) or day surgery patients (under the condition where $B_D \geq 5$ in Table 2). While, under decentralized mode, bed and operation room resources are shared between day surgery patients and elective surgery patients, so it is more equitable. Under centralized management mode, when allocation is more advantageous to elective surgery patients (i.e. fewer beds are allocated to day surgery patients, and more elective surgery patients are served at the cost of having fewer day surgery patients served than decentralized mode), fewer patients are served within equal time period and the average waiting time before admission is much longer when compared to decentralized mode. When allocation is more advantageous to day surgery patients (i.e. more beds are allocated to day surgery patients, and more day surgery patients are served at the cost of having fewer elective surgery patients served than decentralized mode), a lot more patients are served within equal time period and the average waiting time before admission is much shorter when compared to decentralized mode. While the development of day surgery is encouraged, centralized management mode is a better choice. However, if bed allocation is far from reasonable under centralized management mode, it will be exceedingly advantageous to one of the two types of patients, such as the result in Table 2 when $B_D < 4$ or $B_D > 5$. For the convenience of applying the predictions in hospital management, the following work in this paper use simple calculation to explore the beds number allocate to day surgery patients under CECD, i.e. situation in Table 2 where $B_D = 5$, that is advantageous to day surgery and has equity closest to decentralized mode. Taking no account of limitation to operation room resource, we conduct a large number of simulation experiments in this paper and obtain the data of beds number under CECD with distinct arrival and service time. By screening out characteristic variables, we predict the beds number under CECD through linear regression and neural network in the following sections.

3 Linear Regression Prediction

In this section, we use linear regression to predict the beds number allocate to day surgery patients under CECD. Table 3 shows the value of parameters used in regression, where input parameters are shown in the first 11 columns, MV_{OT_D} denotes the mean value of OT_D, SD_{OT_D} denotes the SD of OT_D and so on. B_{CECD} denotes the output of predictions. 100 sets of experiment data are used.

Table 3. Value of input and output parameters used in regression

λ_D	λ_E	MV_{OT_D}	SD_{OT_D}	MV_{OT_E}	SD_{OT_E}	MV_{RT_D}	SD_{RT_D}	MV_{RT_E}	SD_{RT_E}	B	B_{CECD}
10	13	45	21	130	55	892	430	3470	1626	26	4
25	9	95	13	125	43	1098	132	1427	1046	24	13
13	16	44	25	165	58	744	451	1645	2394	11	2
...	...	...	...	...	...	...	...	...	...	...	...
26	11	53	17	78	22	1059	150	1151	949	18	9
25	11	55	26	67	54	1161	397	3167	1825	27	10

Let all the input and output in Table 3 stay unchanged, we enter them into regression equation and find redundant variables in the input parameters. Through stepwise regression, we identify MV_{OT_D}, SD_{OT_D}, MV_{OT_E}, SD_{OT_E}, SD_{RT_D} and SD_{RT_E} as redundant variables. In hospital management practice, the operation time is usually shorten than the length of hospital stay after surgery, so MV_{OT_D}, SD_{OT_D}, MV_{OT_E}, SD_{OT_E} are identified as redundant variables.

We let C_0 to C_5 represents constant, coefficient of λ_D, coefficient of λ_E, coefficient of MV_{RT_D}, coefficient of MV_{RT_E} and coefficient of B. The results of linear regression are shown as Tables 4 and 5.

Table 4. Coefficient of linear regression of the input

C_0	C_1	C_2	C_3	C_4	C_5
0.5248	0.2386	-0.4378	0.0036	-0.0008	0.2726

Table 5. Test indicators of linear regression

R^2 statistic	F statistic	p value	An estimate of the error variance
0.9162	205.5406	0.0000	1.0990

As can be seen from Table 5, the regression equation is valid. So B_{CECD}, the beds number allocate to day surgery patients under CECD can be represented by the formula (1).

$$\begin{aligned} B_{CECD} = 0.5248 + 0.2386\lambda_D - 0.4378\lambda_E + 0.0036MV_{RT_D} \\ - 0.0008MV_{RT_E} + 0.2726B. \end{aligned} \tag{1}$$

The actual beds number allocate to day surgery patients under CECD is integer, thus, we rewrite Eq. (1) by rounding it to integer and we obtain it denoted as ${B_{CECD}}^*$ by the formula (2).

$$\begin{aligned} B^*_{\text{CECD}} = \lfloor 0.5248 + 0.2386\lambda_D - 0.4378\lambda_E + 0.0036MV_{RT_D} \\ - 0.0008MV_{RT_E} + 0.2726B + 0.5\rfloor. \end{aligned} \tag{2}$$

The actual and the predicted beds number allocate to day surgery patients under CECD calculated according to formula (2) are shown in Figs. 1 and 2:

The pentagrams in Fig. 2 denote entirely accurate predictions. From Figs. 1 and 2, we can see that not much predictions are entirely accurate, but the maximum error is 3. The results of the key parameters which reflect regression effect are: $R^2 = 0.9162$, $F = 205.5406$, $p = 0.0000$. Parameter $R^2 = 0.9162$ indicates

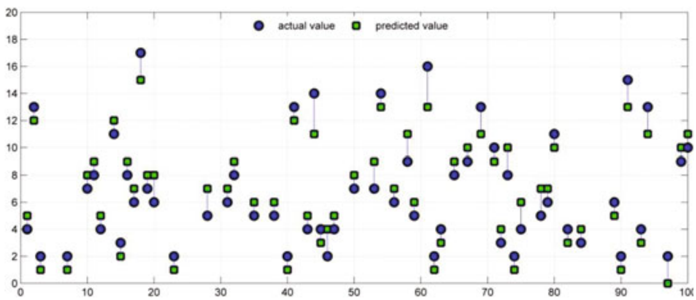


Fig. 1. The predictions with fluctuation using linear regression

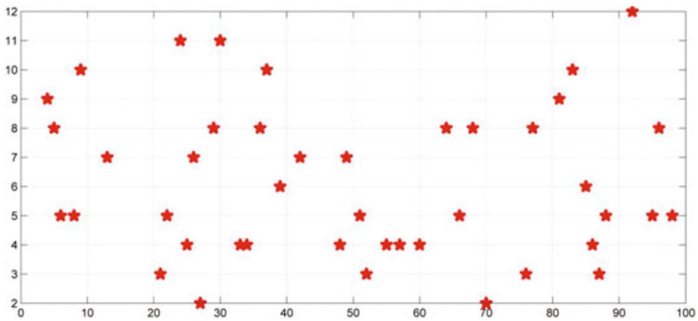


Fig. 2. The entirely accurate predictions using linear regression

that the fitting result of the model is satisfied, parameter $p < \alpha$, which indicates that selection of every variable in the regression equation is significant. The results of Table 5 indicate that Eq. (1) has a satisfied fitting result. The regression equation can provide decision supports for beds allocation under centralized management mode.

Next section, we predict the number of B^{*}_{CECD} with neural network method.

4 Neural Network Prediction

It has identified the input layer has five units. And there is one output unit represents the beds number allocate to day surgery patients under CECD. Neural network theory has been proved that if the number of hidden layer units can be set free, then with three's like I/O node characteristics can approximate any continuous function with any precision. Therefore, we choose networks with three architectures, that is only one hidden layer. And there are eight units at hidden layer.

The error trend after we trained 500 times are seen as Fig. 3.

Based on the neural network above, the results of prediction we obtained are as follows (Tables 6, 7 and 8).

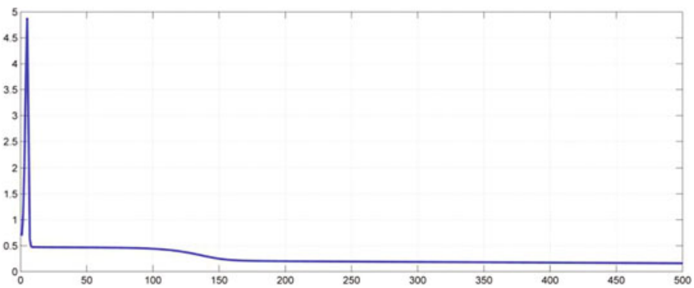


Fig. 3. The error trend using Neural Network Prediction

Table 6. The value of $w1$

$B1$	$B2$	$B7$	$B9$	$B11$
−0.4018	0.3646	−0.2478	0.1699	−0.1319
−0.2532	0.2661	−0.0389	0.067	−0.3323
−0.2306	0.2747	−0.0048	0.0475	−0.3204
−0.1817	0.2614	0.0323	0.0675	−0.3431
−0.353	0.3008	−0.1178	0.0916	−0.2154
0.6947	−0.9056	0.4409	−0.2699	0.7933
−0.3098	0.2657	−0.0423	0.0847	−0.2985
0.0418	0.0656	0.1519	0.0017	−0.1027

Table 7. The value of $w2$

−1.2730	−0.9122	−0.7375	−0.6258	−1.1009	1.2214	−0.9976	−0.0026

Table 8. Threshold of hidden layer and output layer

$B1$								$B2$
−1.0442	−1.6046	−1.7828	−1.8679	−1.3586	−1.7432	−1.4882	−2.0515	0.5287

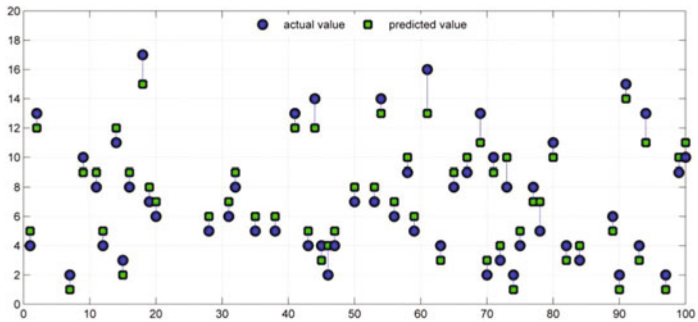


Fig. 4. The predictions with fluctuation using Neural Network Prediction

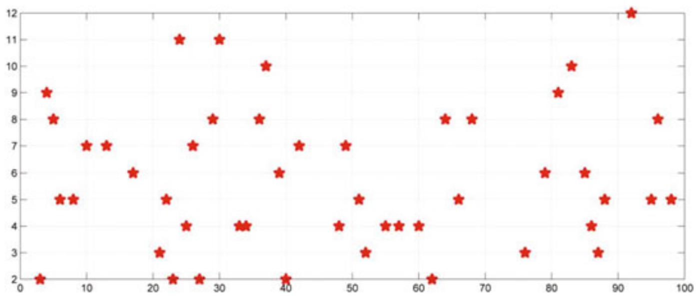


Fig. 5. The entirely accurate predictions using Neural Network Prediction

The pentagrams in Fig. 5 denote entirely accurate predictions. From Fig. 4 and Fig. 5, we can see that not much predictions are entirely accurate, but the maximum error is 3. And the neural network above can provide decision supports for beds allocation under centralized management mode.

5 Conclusion

In this paper, we firstly conduct a large number of simulation experiment using Matlab and prove that centralized management mode is more advantageous to either elective surgery patients or day surgery patients, it depends on the allocation of beds; while decentralized management mode is the most equitable. Then, based on the large amount of simulation data, we screen out characteristic variables and obtain the regression equation of the CECD beds number, the general linear regression equation is rounded to integer as

$$\begin{aligned} B^*_{\mathrm{CECD}} = \lfloor & 0.5248 + 0.2386\lambda_D - 0.4378\lambda_E + 0.0036MV_{RT_D} \\ & - 0.0008MV_{RT_E} + 0.2726B + 0.5\rfloor. \end{aligned}$$

The regression equation satisfied fitting results, the maximum error is 3, and only a few recorded errors are 3, most are 1 and 2 besides the accurate predictions. And the beds prediction equation in this paper, which is simple and convenient for calculation. Then we use artificial neural network to do the same prediction. And the results are similar to linear regression. The maximum error is 3, and only a few recorded errors are 3, most are 1 and 2 besides the accurate predictions. Both linear regression equation and the neural network above can provide decision supports for beds allocation under centralized management mode.

The deficiency in this paper is that we only consider beds number without considering the limitation to operation room resource. We will take account of the limitation to operation room resource to enrich our study next. Moreover, we will figure out other methods to improve accuracy of the prediction.

Acknowledgement. This paper is supported partly by the National Natural Science Foundation of China under Grant Nos. 71532007, 71131006 and 71172197.

References

1. Abdullah L (2014) Modeling of health related quality of life using an integrated fuzzy inference system and linear regression. Procedia Comput Sci 42:99–105 (in Chinese)
2. Aguiar F, Torres R et al (2016) Development of two artificial neural network models to support the diagnosis of pulmonary tuberculosis in hospitalized patients in Rio de Janeiro, Brazil. Med Biol Eng Comput 54:1–9
3. Bonellie S (2012) Use of multiple linear regression and logistic regression models to investigate changes in birth weight for term singleton infants in Scotland. J Clin Nurs 21:92–114
4. Gholipour C, Rahim F et al (2015) Using an artificial neural networks (anns) model for prediction of intensive care unit (icu) outcome and length of stay at hospital in traumatic patients. J Clin Diagn Res 9:1096–1105
5. Launay C, Rivière H et al (2015) Predicting prolonged length of hospital stay in older emergency department users: use of a novel analysis method, the artificial neural network. Eur J Intern Med 26:478–482 (in Chinese)
6. Luo L, Luo Y et al (2014) Difference analysis of day surgery's and elective surgery's duration associated with surgery. Stat Inf Forum 29:104–107 (in Chinese)
7. Menke N, Caputo N et al (2014) A retrospective analysis of the utility of an artificial neural network to predict ed volume. Am J Emerg Med 32:614–617

8. Russell R, Malik R et al (2012) Improved estimates of visual field progression using bayesian linear regression to integrate structural information in patients with ocular hypertension. Invest Ophthalmol Vis Sci 53:2760–2769
9. Schmidt M, Lipson H (2009) Distilling free-form natural laws from experimental data. Science 324:81–85
10. Shakerkhatibi M, Dianat I et al (2015) Air pollution and hospital admissions for cardiorespiratory diseases in Iran: artificial neural network versus conditional logistic regression. Int J Environ Sci Technol 12:504–523
11. Shi H, Chang H et al (2013) Breast cancer surgery volume-cost associations: hierarchical linear regression and propensity score matching analysis in a nationwide taiwan population. Surg Oncol 22:178–183 (in Chinese)
12. Tsai P, Chen P et al (2016) Length of hospital stay prediction at the admission stage for cardiology patients using artificial neural network. J Healthc Eng 6:5–8 (in Chinese)
13. Waltz D, Buchanan B (2009) Computer science. Automating science. Science 324:43–44
14. Wise E, Hocking K, Brophy C (2015) Prediction of in-hospital mortality after ruptured abdominal aortic aneurysm repair using an artificial neural network. J Vasc Surg 62:8–15 (in Chinese)
15. Xian L, Feng C (2015) Predicting menopausal symptoms with artificial neural network. Exp Syst Appl 42:8698–8706 (in Chinese)

A Standardization Methodology for Visual Management in Lean Supply Chain Environments

André Simas(✉) and Virgilio Cruz-Machado

UNIDEMI (Department of Mechanical and Industrial Engineering), Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
af.simas@campus.fct.unl.pt

Abstract. A rather important tool of Lean used in a plant is Visual Management, because people pay more attention to everything that is visual than any other sense. This kind of eyes management allows the abnormalities to stand out and simplify the work that needs to be done. Visual Management within an organization, assumes that all the tools used in the plant are standard, to be the same in the rest of the organization. In order to help organizations about the develop of visual standards, the author proposed to develop a Standardization Methodology divided in two sections, the Current Situation Analysis Methodology and a Future Design Methodology. The first one refers to all stages that precede the analysis of instruments of some plants, while the second one concerns the creation of improvements. The enhancement proposals were focused on processes that could be an advantage to the existence of a standard tool. Most of these tools are low cost of employment and there are some advantages in their use: they allow more transparency on processes and they make the shop floors look like cities.

Keywords: Lean · Visual management · Standardization · Methodology

1 Introduction

Nowadays, the consequences of the Great Recession are still quite noticeable, a downfall that started in 2008 with the Lehman Brothers Holding Inc., an investment bank, declaring bankruptcy. Several other banking institutions followed the same path after the beginning of crisis. After that period, the companies focused on a clear objective, use innovation to be more competitive.

The automotive industry has also suffered some financial distress, leading manufacturing organizations to start to think of new production methods and a better application of the tools used at the shop floor. This allows companies to pursue operational excellence with an improvement of their performance,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_95

providing high quality products earlier in order to reduce costs [2]. Worldwide, the companies start to encourage a Lean culture on their employees due to the competitive nature of this industry, caused by emerging markets which produce similar products at lower prices. To face those markets, the automotive industry must be constantly updated, developing new products to customers.

Lean Thinking is the basis of implementation within companies, providing a guide on how they can do more with less, less effort, equipment, time and space. This can be divided in a set of five principles: value, value stream, flow, pull and perfection [10].

Implementing Lean can be a difficult process, but when it is well applied becomes an essential characteristic within an organization, playing a crucial role leading up to the intended results. Some authors said this philosophy is determinant for the survival and constant innovation of organizations. One tool from Lean culture is Visual Management, an important tool used at the shop floor, which is defined as signs and other forms of visual information used to simplify the workplace and make it easy to recognize abnormalities on processes [4].

2 Visual Management

Visual Management is a Lean tool that has the objective of giving visual information or displaying requirements to set directions [3]. This tool is frequently used at manufacturing industries, but has currently expanded to other business industries [1]. This concept was created to highlight abnormalities directly in the workplace, thus helping operations and processes as soon as a problem occurs [8]. Giving the right information to employees at the right moment is a vital characteristic to improve the performance of a company [3]. That kind of information can be provided by signs, labels or a colour code. The use of this type of information eliminates the "guess", searching and a cumulative of information or material [6].

Many authors, such as Wilson [9] don't use the term "visual management" but rather transparency, because this tool allows the observation of processes on time. An operator can see what is happening in a process and he can change or adjust it if an abnormality is to occur.

Within Visual Management there are tools that support operators' tasks and highlight any abnormalities on processes. There are two different sets of tools [3]:

- Understanding processes: tools mainly focused on visualising the information to allow a better understanding of the processes. Ex.: value stream mapping, flow charts, A3 and area name boards.
- Performance of processes: tools that are used to communicate requirements and manage the efficiency of the processes. Ex.: andon lights and boards, Kanban and KPIs screens.

Many of these signals were created to control robust processes, allowing to draw attention to many processes at once and act immediately [7].

Some authors argue that Visual Management is the basis of sustainable improvement, promoting the contribution of all operators in management activities and a quality enhancement of processes. This can only be obtained when the information is conveyed in a succinct and clear manner [5].

The use of Visual Management has many benefits:

- less time spent understanding the information;
- visible abnormalities, installation of devices/signals;
- speed, the problems are eliminated earlier;
- involvement of all, promoting a continuous improvement;
- standardization, processes up to date.

On this list, only some of the most important benefits were referred, there are many others that are related with this Lean technique. For Parry & Turner [8] visual management must be simplified, only information that adds value to the processes should be displayed.

3 An Approach to Establish a Standardization of Visual Tools

Mostly organizations that have visual management on the mindset of their employees and use it constantly, can make the process flow through these visual aids. Employees must interact with the tools from visual management to continue the process, if they don't have that support the process doesn't flow smoothly until the end.

Every workplace has visual aids. This aids can be presented in many shapes, for example: instructions of how to do a stage of process; directional signals; photographs and boards. However, this information is often neglected because many organizations don't follow the requirements of visual management. These requirements are all information to bring added value, easily understandable and must be placed at visible places and updated.

An organization that defines very good standards for visual controls has more benefits and employee efficiency. Then a methodology will be described, that provides a simple guide for organizations who want to standardise their visual tools.

3.1 Methodologies

A methodology characterizes all the stages to follow from a process. It is a detailed explanation of all actions of the work until the expected result is reached.

For better support, a Standardization Methodology divided in two sections was developed:

- Current Situation Analysis Methodology - this refers to all stages that precede the analysis of instruments of some plants;
- Future Design Methodology - this refers to the creation of improvements.

All stages must be strictly followed for the success of a project, if the implementer doesn't respect the stages, the results can't be truly trusted. This Methodology was designed for organizations that have multiple plants worldwide and the tools are often unequal, but can be used in all organizations with few adjustments.

(1) Current Situation Analysis Methodology

This Methodology is the initial phase of a standardization, representing the analysis of information that the implementer collects from an organization (Fig. 1).

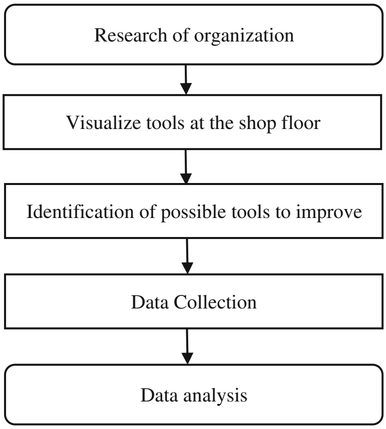


Fig. 1. Stages of current analysis methodology

① Research of organization

This phase must be the most detailed one considering the resources available, documents or procedures that are used by the organization in the composition of the workstation, which incorporates the visual controls. The person who wants to implement this methodology should know if the organization has some Lean policy that can match with this methodology.

② Visualize tools at the shop floor

The objective of this stage is to get used to the tools that are being used at the shop floor, through its visualization in operation. This knowledge is the transition from theory, studied in the previous phase, to practice. The need of seeing tools at the shop floor can clarify some misunderstandings of theory information from documents and can be a way to identify abnormalities on tool applications.

③ Identification of possible tools to improve

After observing the tools in operation, the implementer should use critical sense to identify possible tools for improvement. It is important to keep in mind the limitations of the organization, starting with a few number of improvements is better than start with a huge change on tools. Employees tend to accept changes if they are implemented slowly.

④ Data Collection

This is the critical phase of all the methodology, because at this stage is where the implementer gathers the information from plants related to the tools for improvement. This data collection combines interviews, by email or telephone, with photos to support what they said, and visits to other plants from the organization to check on how they apply the tools. The characteristics required from plants are based on how the tool is used, its appearance, to whom it is intended and where it is used inside the shop floor.

⑤ Data analysis

After the data collection is concluded, the information of all plants must be placed on paper at the beginning (like a draft). When the draft is completed, the implementer can introduce that in a computer and print it. It is recommended to print in large formats for easier comprehension. The disposal of information provides a good view of how tools are being used, if similar, or not, to the other plants. When a plant has a good practice of some tool, it must be used in total or partially for improvement of that tool.

(2) Future Design Methodology

This part of the Methodology is the development phase, which include the conception and the implementation at the shop floor (Fig. 2).

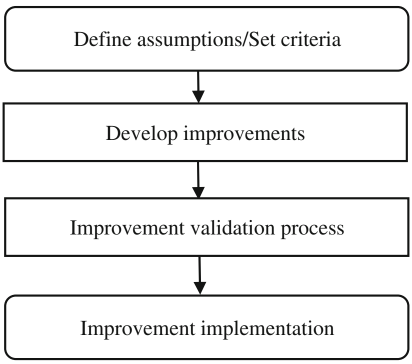


Fig. 2. Stages of future design methodology

① Define assumptions/Set criteria

The conception of improvements is a subject that has some rules and must be studied before presenting the improvements. If there aren't rules or a previous study those tools can be worse than the actual. For an initial phase, it should be guaranteed that some criteria are settled by the organization that uses this methodology. The criteria are:

- Objectivity: must be clear, to not confuse employees;
- Similar: if a plant has a good tool, that tool should suffer a few adjustments to be similar;
- Reliable: improvements proposals shall consider their economic reliability;
- Simple and intuitive: visual tools are easier to retain;
- Respect layouts and colours: when developing improvements, the rules established by organization must be respected.

The need to establish criteria is the basis for a successful implementation of a strategy and for a best probability of being validated by the Central Team.

② Develop improvements

Through analysis of collected information from plants, a brainstorming must be done for chosen tools, following the previously explained aspects. That brainstorming should be done with a team of 2 or more people. A team with more people can obtain multiple outputs, hence different tools. The improvements derive, more or less from the previous presentation, depending on the good level of effectiveness of the tool. The team presents all ideas for improvements. It is recommended to have more than one solution, but it isn't required.

③ Improvement validation process

This phase occurs when the team presents the improvements to the Central Team (team responsible for the application of Lean at plants, with the objective of choosing the best ones to apply). The validation process is the approval and standardization within an organization. A tool is accepted when the proposal shows a better potential than the tool being currently used and will bring more benefits. This comparison is made using a table with advantages and disadvantages of the old tools and the new ones. The Table 1 must fit the following layout:

Table 1. Table template for comparison of tools

Characteristics	New tool	Tool plant 1	Tool plant 2
Characteristic 1			
Characteristic 2			

There are no predefined characteristics, the organization must adapt that subject according to the nature of the tool. To fulfil the Table 1 above the organization must mark with "✓" if the tool has the characteristic or "♣" if not. Through this analysis it's possible to verify if the new tool has, in theory, more potential than the implemented tool.

④ Improvement implementation

After being approved by the Central Team, it's the right moment to start standardizing on all plants. It's a slow process because implementations take time, so it's recommended to act as quickly as possible. The output resulting from the application of this methodology is the implementation of the tools that were subject to improvements at the shop floor. This is the last step of information collection and analysis regarding plants from different regions.

3.2 Application of Methodologies

This subchapter will give the reader a better comprehension of this methodology applied in an organization from the automotive industry. Two topics will be approached where the methodology can be applied, for example, Stopping Points, a subject from routing, and the color code from plants.

Every plant has routes and this routes need stopping points, a specific location near the workplace where the tuggers' drivers do the collection or replacement of materials during the route. After getting the information from plants, through photos of the application of the tools at plants, it was verified that the tool's presentation is different in all plants within an organization. Figures 3, 4, 5, and 6 show the differences in stopping points between some plants.

Fig. 3. Stopping point plant 1

Fig. 4. Stopping point plant 2

Fig. 5. Stopping point plant 3

Fig. 6. Stopping point plant 4

The existence of several ways of presenting the stop signs shows the need to standardize this tool, so every plant can use the same tool.

Another concept that many plants have huge differences between them is the color code, normally used in floor marking. This tool is important to have for a well-organized workplace so it can prevent accidents and injuries. Many organizations use these markings with the intention of enhancing the visual management of the organization. To obtain the data from the plants interviews by phone were held, with photos to support what has been told on interviews. These photos in Figs. 7 and 8 show that there are enormous differences between the color code from each plant.

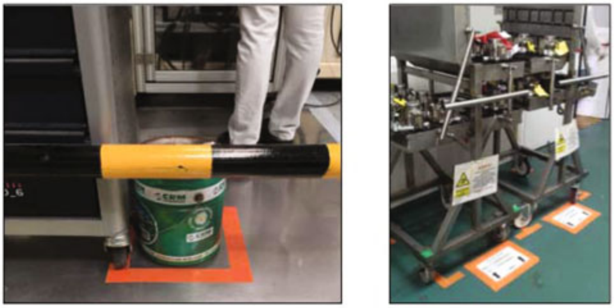


Fig. 7. Orange color for dangerous material [plant 1] (at left) and rework [plant 2] (at right)

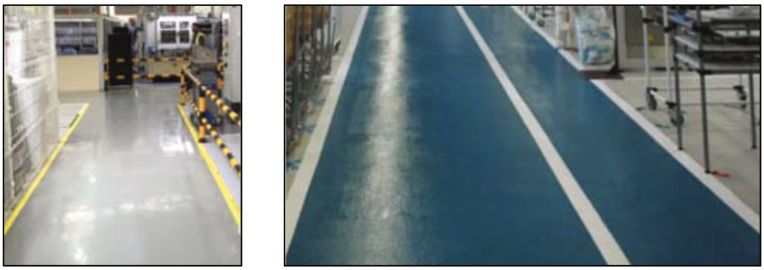


Fig. 8. Differences between two plants at aisles markings

When all information was collected, it was observed that there wasn't a common color code for all plants, every plant has their own code.

3.3 Results

Now, the improvements for the tools mentioned above will be described, including Stop Points and Color Code.

The appearance of Stop Points is a very useful visual tool in routing, they let the driver of the tugger know which route belongs to what and the number of the stop (usually are numbered with 1, the nearest stop from warehouse and so on). The decision was to design an easily visible tool with all the necessary information, so the driver doesn't have issues performing the route. The new tool is shown in the Fig. 9.

Fig. 9. New stop sign developed

In this case, this tool has the name of the route it belongs to, a frame with the color of the route and the number of the stop, the other aspects are only aesthetics. For the design of this tool, the criteria established by the methodology were considered. To compare this tool with the ones already implemented, the methodology requests a table with the comparison:

Table 2. Comparison of stop points

Characteristics	New tool	Tool plant 1	Tool plant 2	Tool plant 3	Tool plant 4
Fast interpretation	♡	♡	♣	♣	♣
Visible	♡	♡	♣	♡	♡
Low cost	♡	♣	♡	♣	♡

This tool can be quite financially viable. It is made with paper (A4 format), with the colors from the respective routes allowing a fast interpretation for the tugger's driver (Table 2).

As I said before, the color code is essential to mark workplaces and pedestrian access areas. Those marks should be visible and noticeable by operators, employees and visitors that walk at the shop floor. If a workplace has a tidy

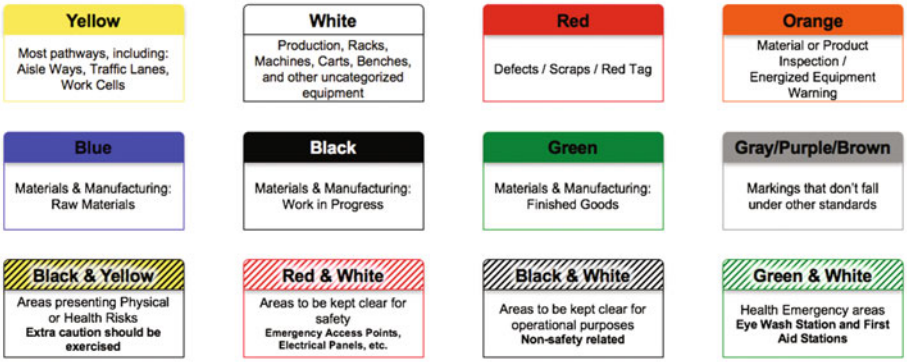


Fig. 10. New color code developed

working environment that stimulates a good performance of an operator. So, a color code was designed to be adopted by all plants within an organization:

From Fig. 10, the objective is the adoption by all plants of a single standardized color code in the organization. For the design of the code, the criteria established by the methodology were considered and was decided not to do the comparative table. Despite the differences, there is no comparison term for the color code. The colors present in the code are the minimum required for a plant, however, if any plant needs more combinations of colors they can define them.

4 Conclusions

Standardization of visual management within an organization, with a lot of plants, is a complex and long journey process. You need to transform an entire organization, from tools that are being standardized to the mindset of employees, who will contact with the new tools. The tools to improve must be requested as they come across a tool that can be improved.

The benefits of having all visual management standardized are huge: reduction of wasting time using confuse tools; an increased efficiency of operators, tools are designed for them; simplifies the work environment and the seek for excellence. Normally the output (improvement tools) manages to convey the basic principles of visual management, through "eye management", always prevailing simplicity, transparency, and clarity.

References

1. Bateman N, Philp L, Warrender H (2016) Visual management and shop floor teams - development, implementation and use. Int J Prod Res 7543:1–14
2. Belekoukias I, Garzareyes JA, Kumar V (2014) The impact of lean methods and tools on the operational performance of manufacturing organisations. Int J Prod Res 52(18):5346–5366

3. Eaidgah Y, Maki AA et al (2016) Visual management, performance management and continuous improvement: a lean manufacturing approach. Int J Lean Six Sigma 7(2):187–210
4. Emiliani ML, Stec DJ (2005) Leaders lost in transformation. Leadersh Organ Dev J 26(5):370–387
5. Hines PA, Found P et al (2008) Staying Lean: Thriving, Not Just Surviving. Society for American Archaeology
6. Machado VC, Leitner U (2010) Lean tools and lean transformation process in health care. Int J Manag Sc Eng Manag 5(5):383–392
7. Ortiz CA, Park M (2010) Visual Controls: Applying Visual Management to the Factory. Productivity Press, New York
8. Turner GC, Parry CE (2006) Application of lean visual process management tools. Prod Plann Control 17(1):77–86
9. Wilson L (2010) How to Implement Lean Manufacturing. McGraw-Hill, New York
10. Womack J, Jones D (2003) Lean Thinking, Banish Waste and Create Wealth in Your Corporation. Free Press, New York

Resource Optimization Management

Online Fault Detection in Solar Plants Using a Wireless Radiometer in Unmanned Aerial Vehicles

Carlos Quiterio Gómez Muñoz(✉), Alfredo Peinado Gonzalo, Isaac Segovia Ramirez, and Fausto Pedro García Márquez

Ingenium Research Group, Castilla-La Mancha University, Ciudad Real, Spain
CarlosQuiterio.Gomez@uclm.es

Abstract. A novel Non-Destructive Test (NDT) is presented in this paper. It employs a radiometric sensor that measures the infrared emissivity of the solar panel surface embedded in an unmanned aerial vehicle. The measurements provided by the sensor will determine if the panel is healthy, damaged or dirty. A thermographic camera has been used to check the temperature variations and validate the results by the sensor. The study shows that the amount of dirt influences the temperature on the surface and the energy generated. Similarly, faults in photovoltaic cells influence the temperature of the panel. The NDT system is less expensive than traditional thermographic sensors or cameras. Early detection of these problems, together with an optimal maintenance strategy, allows to reduce costs and increase the competitiveness of this renewable energy source.

Keywords: Infrared sensor · Radiometry · Solar plants · Photovoltaic panels · Fault detection and diagnosis

1 Introduction

1.1 Renewable Energies

The whole energy produced in the Earth has been obtained, directly or indirectly, from the Sun. Even fossil fuels, which come from the decomposition of plants that have previously needed the sun to perform photosynthesis. These fossil fuels are oil, coal and natural gas, and they are in limited reserves.

From the industrial revolution, fossil fuels consumption has increased without considering the environmental effects, such as global warming or climate change. However, as they are finite resources, alternative forms of fossil fuels must be sought to obtain energy and to avoid the pollution they produce. These measures are being carried out globally, led by the Intergovernmental Panel on Climate Change (IPCC).

The renewable energies are being developed and employed to use a clear production, e.g. in Spain they produce 28.6% of the energy consumed [9,16]. Renewable energies are those based on the use of virtually endless resources, e.g. the sun (Fig. 1), wind or ocean currents, or resources that are renewed periodically if they are used sustainably, such as biomass or biofuels. In general, its production is more expensive than

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_96

Fig. 1. China and Pakistan project for Solar plant Solar Quaid Azam installed in Pakistan (http://www.qasolar.com/)

fossil fuels, but if does not generate polluting emissions during their production. These include: Wind energy [14,22], tidal, solar (Fig. 1), biomass and biofuels, geothermal and hydraulic.

1.2 Non-Destructive Experimenting

Non-Destructive Experimenting (NDT) is a type of experiment that is performed on parts and pieces to verify its state and operation without interfering in the integrity of the material or its usefulness to the service [6,10,19,24]. They are carried out in recognition and preventive maintenance work in all types of industrial areas. These experiments are used to detect internal and surface defects [8,18,20,21]. Some of these experiments are: Radiography with X-ray and Gamma, magnetic particles, Eddy currents, ultrasound experimenting [7,11,17] and infrared thermography [1]. Infrared thermography consists in checking the state of a part or a structure from a thermal image. They are used in many industries and they can be applied in any production phase.

2 Infrared Thermography

2.1 Radiation

There are only three methods of heat transmission in physical terms: conduction, convection and radiation. In the case of radiation, a transmission medium is not necessary, the energy is transmitted due to the temperature of the emitter body. Depending on the temperature, the radiation emitted changes. The law of Stefan-Boltzmann, Eq. (1), is used to quantify it. It must be adjusted for real bodies, by adding the emitter parameter to the equation.

$$E_N(T) = \sigma \cdot T^4, \tag{1}$$

where E_N is the emittance of the body, is the Stefan Boltzmann constant, value 5,6710 B -8 W/m^{2}K^4)) and T the body temperature. This emittance is for of a black body, which must be corrected for real bodies through the emissivity. The Planck law is given by Eq. (2).

$$E_N^{\lambda}(\lambda, T) = \frac{C_1}{\lambda^5 (e^{\frac{C_2}{\lambda T}} - 1)}, \tag{2}$$

where E_N is the spectral emitted, C_1 and C_2 are constants, λ the wavelength and T the body temperature. Eq. (2) shows that most of the radiation emitted by bodies at low temperatures is found in the infrared region [15].

Emittance is the proportion of energy that a body can emit relative to the total. In a black body (maximum emitter) is the unit. In any other body, this magnitude is smaller and depends on the temperature. In addition, the characteristic wavelength of the bodies also depends on their temperatures, generating the electromagnetic radiation spectrum.

The thermal behavior of a real body is different from a black body. While the black body absorbs all the energy that reaches its surface and emits throughout the radiation region equally, a real body does not absorb all the energy and does not emit in the entire electromagnetic spectrum.

A real body does not absorb all the radiation, but part of it passes through it or reflects on its surface. In addition, this behavior also depends on the wavelength being studied, that depends the concept of emittance, given by Eq. 3. Emittance is the ratio of energy that a body emits with respect to the total that it can emit with a certain temperature and wavelength [15].

$$\varepsilon^{\theta,\varphi,\lambda}(\theta, \varphi, \lambda, T) = \frac{E^{\theta,\varphi,\lambda}(\theta, \varphi, \lambda, T)}{E_N^{\theta,\varphi,\lambda}(\theta, \lambda, T)} = \frac{I^{\theta,\varphi,\lambda}(\theta, \varphi, \lambda, T)}{I_{E,N}^{\theta,\varphi,\lambda}(\lambda, T)} \tag{3}$$

being ε the emissivity, θ and φ emission angles, E emittance and I the radiation intensity of the body. For this research work, infrared radiation is the most relevant, since it is the dominant thermal radiation in this temperature range. Therefore, the camera and the sensor are designed to collect this type of radiation.

2.2 Applications

The main function in this paper of infrared thermography is the visualization of elements whose surface temperature variations show their state [2,3,12]. Among its applications include the following:

(1) Electrical and mechanical systems: It is possible to observe the hot spots of the systems and areas subjected to higher thermal stresses by comparison on the surface of the system [13].
(2) Structures: Any irregularity in the temperature of the facade can correspond to different conditions in buildings, such as poor insulation, the emittance of an area with a greater amount of bricks than concrete, or even an insect nest within the facade.
(3) Welds: Temperature variations in a pipe are usually due to areas where welds have been made. If these variations show a very irregular appearance, it is probably due to a bad union
(4) Surveillance: Control of emissions, traffic, fire prevention, etc.
(5) Medicine: Monitoring of diseases and anomalies in the human body.

The infrared thermography is done by a non-radiometric camera to check the results obtained with the sensors.

3 Experimental Platform

3.1 Radiometer

They are used to check the temperature and the radiance of the measured targets. Both sensors are used to contrast the data obtained with each one in the different case studies.
(1) Campbell SI-111 Sensor

This sensor is connected to an Arduino board, which processes the voltage data and process that information to obtain the temperature in Celsius degrees. The system is capable of making accurate measurements to the bodies that are within its field of vision. The SI-111 Precision Infrared Radiometer features a thermistor to measure its internal temperature, and a thermopile for the target temperature using a germanium lens.

Using Stefan-Boltzmann's law of radiation, and the voltages in the thermopile and the thermistor, the target temperature is calculated [23].

The temperature measured by the sensor is the average of the temperatures of its view range, receiving between 95 and 98% of the infrared radiation of the view range and between 2 and 5% from outside. Figure 2 shows the view range of the SI-111 sensor with half-angle of 22.5o. The view range depends of the distance to the area and the inclination with respect to this area (see Fig. 2).

This sensor has been chosen because of its high accuracy, reaction rate to changes in temperature and the large range of temperatures that can operate. Its size and weight allow a great manageability. Its installation on the unmanned aerial vehicle allows the global system to performance more efficient and flexible inspections.

3.2 Arduino Platform and Wi-Fi Shield

For this research, Arduino Uno G3 motherboard is used and the Shield Wi-Fi extension, which grants wireless features to the original board. This device has been chosen due to its great extension in this type of applications and its low consumption. The Arduino

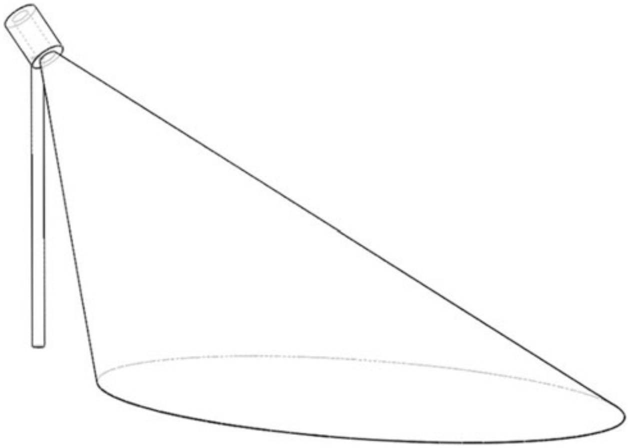


Fig. 2. SI-111 sensor field of view

(a) (b)

Fig. 3. Arduino Uno G3 (a) and Wi-Fi Shield expansion card (b)

Uno R3 (Fig. 3a) is a motherboard that uses the ATmega328 microcontroller and is powered by either USB cable connected to a computer or external power from a battery.

The Wi-Fi Shield component is an Arduino expansion card. It allows the transmission of data via wireless connection by Wi-Fi (Fig. 3b).

Data obtained by the sensor is processed and sent via Wi-Fi to the online platform Thingspeak to monitor the results (Table 1).

3.3 Thermographic Camera

The camera is non radiometric used to check the data obtained with the sensors and to see the temperature differences of the observed pieces. This camera consists in assigning a determined colour to each level of radiation that it receives. These levels are calculated for each resolution pixel that the camera has, creating a thermographic picture (Figs. 4 and 5).

Table 1. Technical characteristics of the Campbell SI-111 sensor

Input voltage	2.5V excitation for the thermistor.
Absolute accuracy	$\pm$ 0.2 °C between [−10 and 60] °C $\pm$ 0.5°C between [−40 and 70] °C
Response time	Less than 1 s for changes in target temperature.
Field of view	22° of half-angle
Operating environment	Highly water resistant designed for continuous outdoor use. Operating range −55 °C to 80 °C 0 to 100% relative humidity

Fig. 4. Thermographic camera Flir VUE

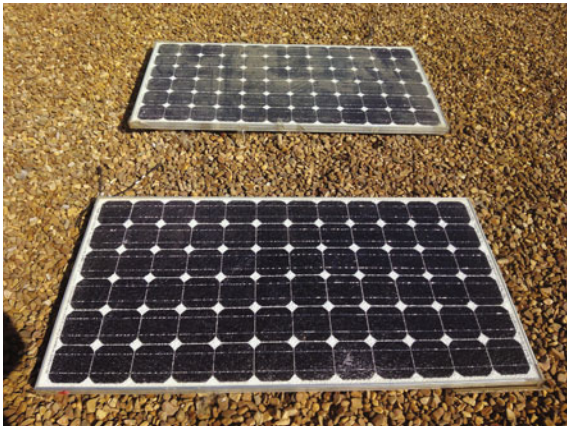


Fig. 5. Solar panels used in the experiments

3.4 Photovoltaic Solar Panels

These panels were used in the experiments to check their temperature and voltage at the terminals. Their technical characteristics are shown in Table 2.

Table 2. Technical characteristics of photovoltaic solar panels

Manufacturer's name	TSM-170D
Max power	2.5 W
Maximum voltage	35.8 V
Maximum intensity	4.76 A
Open circuit voltage	43.6 V
Short circuit current	5.25 A
Maximum system voltage	700 V
Work temperature	[−40 °C,90 °C]
Dimensions	1581 × 809 × 40 mm
Cell type	Monocrystalline
Number and size of cell	72 pieces, 125 × 125 mm
Type of glass and width	3.2 mm tempered
Weight	15.6 kg

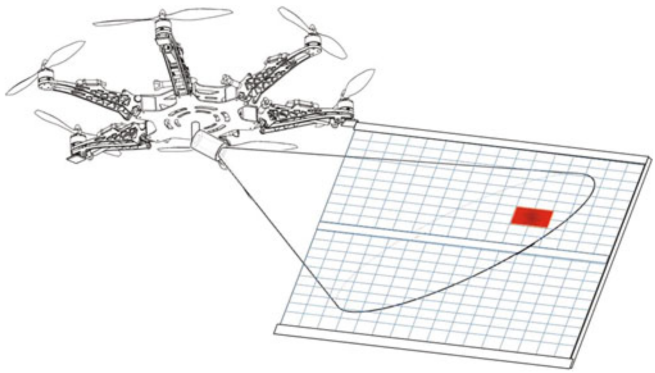


Fig. 6. Representation of the air vehicle taking a measurement with the wireless sensor mounted on the Gimbal mount

3.5 Unmanned Aerial Vehicle DJI Spreading Wings S900

This unmanned aircraft operated by remote control is designed for reconnaissance, monitoring and taking of measurements and images, thanks to its control system that allows a great stability and handling.

The hexacopter can operate around 18 min autonomously, following pre-established routes. It is made of carbon fiber to lighten weight, which allows a maximum load capacity of 8.2 kg. The gimbal is a crucial component for this application for increasing the stability and accuracy of the measurements. This support contains the measurement equipment (sensor, camera, Arduino card and electronic components) in the drone. A schematic representation of the drone collecting data in a photovoltaic plant is shown in Fig. 6.

This system allows to inspect solar plants in a totally automatic and autonomous way. It can inspect large areas, and a very large park can be inspected by sections. The system sends online all the information about the state of the plates (dirt and defects), generating alarms and reports of each module.

4 Case Study

The objective of the experiments was to detect dirt on the solar panels [4]. The experiments were carried out in three different scenarios: In the first scenario, the solar panel was clean; the second one, mud was added in half solar panel; third scenario, the solar panel was totally covered by mud. In each experiment, the temperature values collected by the radiometer and the volts generated by the solar panel were obtained.

The incidence of the sun is crucial in the emissivity values collected by the radiometer. For this reason, they were carried out three cases of study at different hours to study the incidence of the sun in the results. The first case study was performed at 10:00 am. The second case study at 12:00 noon. The last case study was carried out with absence of direct light because of clouds (Figs. 7, 8 and 9).

Fig. 7. Solar panel without dirt on its surface

Fig. 8. Solar panels covered in half of mud

Fig. 9. Solar panels totally covered of dirt

Table 3. Results obtained in the case study 1

	Temperature (°C)	Voltage (V)
Clean panel	46.2	36.2
Partially covered panel	44.7	35.7
Fully covered panel	44.1	35.3

4.1 Case Study 1: Incidence of Sunlight at 10:00 AM

Experiments have been performed at 10:00 am to check the temperature reached by the panels when, due to the incidence of light, the panel does not reach its maximum power generation. Table 3 shows the obtained temperature and generated voltage by the panel.

It is observed an evident difference of temperatures between the clean panel and the panel with dirt on the half of its surface. There are also minor differences between the half dirty panel and the panel totally covered by mud. Figure 10 shows the different temperatures obtained in the series of experiments represented by columns. The ambient temperature series has been added to contrast with those acquired by the panels in each case shown in the form of a line.

These relationships are proportional to the voltages generated by the solar panel. Figure 11 shows the influence of the mud in the voltage at 10:00 am [5].

4.2 Case Study 2: Incidence of Sunlight at 12:00 PM

These experiments have been performed to observe the differences in temperature and energy production of the panels when the panel reach its maximum of production, i.e. when the light strikes perpendicularly on the panels (Table 4).

The temperature variations are bigger than in the first case study, as well as the voltage. This is because of the intensity of the light is much greater and it makes more evident the cases where the panel is covered with dirt. Greater differences in temperature are observed in Fig. 12, as well as higher temperatures.

A greater voltage difference is shown in Fig. 13. It is due to a greater inequality of sunlight incident between the panel without dirt and completely covered.

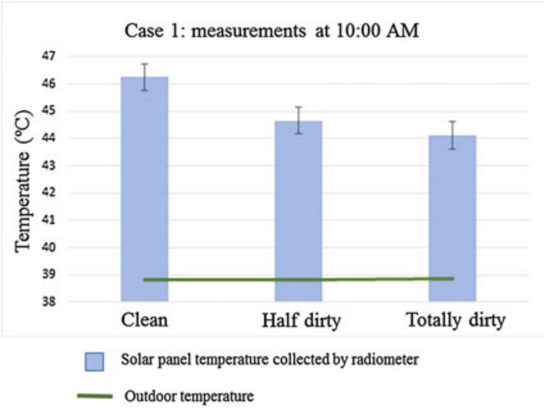


Fig. 10. Comparison between the different panel temperatures in case study 1

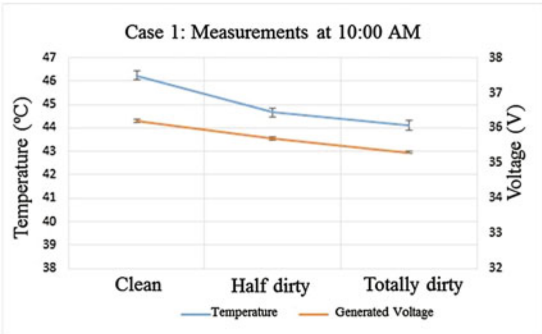


Fig. 11. Comparison between different panel temperatures and voltages in case study 1

Table 4. Results obtained in the second experiments

	Temperature (°C)	Voltage (V)
Clean panel	61.9	37.2
Partially covered panel	57.9	34.4
Fully covered panel	51.8	33.5

4.3 Case Study 3: Absence of Direct Light (18:00 P.M.)

These experiments have been performed at 18:00 pm with clouds to study the effectiveness of the system for the detection of dirt under these conditions. The temperature results obtained are shown in Table 5.

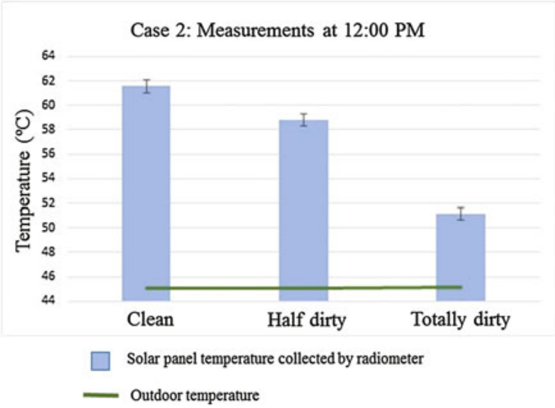


Fig. 12. Comparison between different panel temperatures in case 2

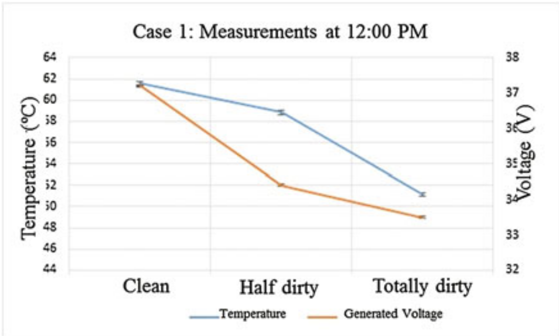


Fig. 13. Comparison between different panel temperatures and voltages in case study 2

Table 5. Measurements obtained in the case study 3

	Temperature (°C)	Voltage (V)
Clean panel	31.4	27.3
Partially covered panel	31.2	27.3
Fully covered panel	30.9	27.3

The differences of temperatures are lower, with a large margin of error that does not allow to know exactly the condition of the panel. Figure 14 shows the low temperature variations between the different experiments.

The voltage in each case are similar for all case studies. In these conditions, it is concluded that the system is not effective to detect the dirt in the solar panels.

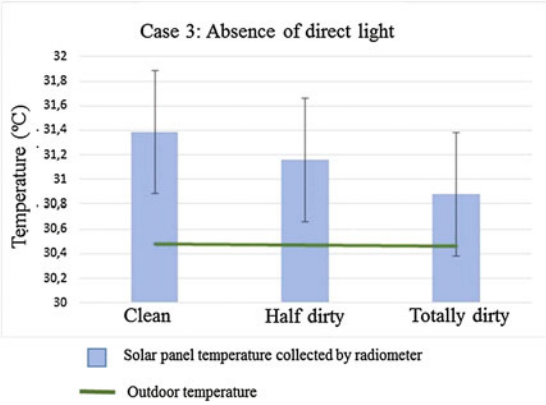


Fig. 14. Comparison between the different panel temperatures in case study 3

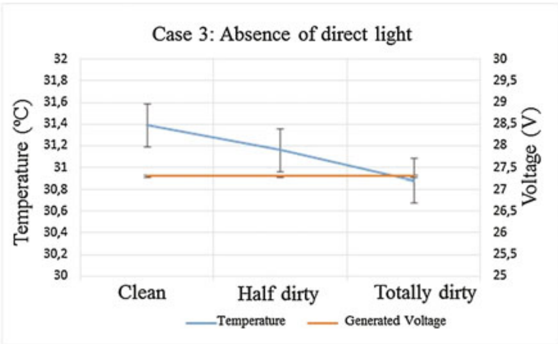


Fig. 15. Comparison between different panel temperatures and voltages in case study 3

5 Conclusions

Dirt and dust in solar panels is a common problem in photovoltaic plants, which are usually located in desert areas with lots of sand and dust. A novel non-destructive test system is proposed in this paper, based on: a radiometer; Arduino; a Wi-Fi Shield, and; a UAV for inspecting the state of the photovoltaic panels.

The low weight and cost of the system can reduce the costs in operation and maintenance of the whole solar plant. The system can be automated, reducing the inspection time and the costs (Fig. 15).

Three different conditions of the solar panel were analysed: Clean, half covered by mud and totally covered. The infrared radiometry of the panels was read and the voltage of the panels. The experiments were performed in three different solar conditions to increase the actuary of the experiments: at 10 am, 12:00 pm and with clouds and without direct sunlight. The absence of direct light does not allow identification among the three cases.

Acknowledgements. The work reported herewith has been financially by the Spanish Ministerio de Economíay Competitividad, under Research Grant Ref.: RTC-2016-5694-3.

References

1. Acciani G, Simione G, Vergura S (2010) Thermographic analysis of photovoltaic panels. In: International conference on renewable energies and power quality (ICREPQ10). Granada, Spain, March, pp 23–25
2. Ancuta F, Cepisca C (2011) Fault analysis possibilities for pv panels. In: International youth conference on energetics, pp 1–5
3. Bazilian MD, Kamalanathan H, Prasad DK (2002) Thermographic analysis of a building integrated photovoltaic system. Renew Energ 26(3):449–461
4. Dorobantu L, Popescu M et al (2011) The effect of surface impurities on photovoltaic panels. In: International conference on renewable energy and power quality
5. Fares Z, Becherif M et al (2013) Infrared thermography study of the temperature effect on the performance of photovoltaic cells and panels. In: Sustainability in energy and buildings, Springer, pp 875–886
6. García Márquez FP, Muñoz G et al (2014) Structural health monitoring for concentrated solar plants. In: 11th International conference on condition monitoring and machinery failure prevention technologie. Manchester, UK
7. Gómez CQ, Villegas MA et al (2015) Big data and web intelligence for condition monitoring: a case study on wind turbines. In: Handbook of research on trends and future directions in big data and web intelligence; information science reference. Hershey
8. Gómez Muñoz CQ, García Márquez FP (2016) A new fault location approach for acoustic emission techniques in wind turbines. Energies 9(1):40
9. Heras-Saizarbitoria I, Cilleruelo E, Zamanillo I (2011) Public acceptance of renewables and the media: an analysis of the spanish pv solar experience. Renew Sustain Energ Rev 15(9):4685–4696
10. Ruiz de la Hermosa González-Carrato R, García Márquez FP et al (2015) Acoustic emission and signal processing for fault detection and location in composite materials. In: Global cleaner production & sustainable consumption conference. Elsevier
11. Jiménez AA, Muñoz CQG et al (2017) Artificial intelligence for concentrated solar plant maintenance management. Springer, Singapore
12. Maldague XP (2002) Thermographic inspection of cracked solar cells. In: Aerosense, p 185
13. Maldague XPV (2002) Introduction to ndt by active infrared thermography. Mater Eval 60(9)
14. Márquez FPG, Tobias AM et al (2012) Condition monitoring of wind turbines: techniques and methods. Renew Energ 46(5):169–178
15. McAdams WH (1958) Heat transmission. McGraw-Hill
16. Montoya FG, Aguilera MJ, Manzano-Agugliaro F (2014) Renewable energy production in spain: a review. Renew Sustain Energ Rev 33(2):509–531
17. Munoz CQG, Arenas JRT, Marquez FPG (2014) A novel approach to fault detection and diagnosis on wind turbines. Glob Int J 16(6):1029–1037
18. Muñoz CQG, Marquez FPG et al (2015) A new condition monitoring approach for maintenance management in concentrate solar plants. Springer, Heidelberg
19. Muñoz CQG, Márquez FPG, Tomás JMS (2016) Ice detection using thermal infrared radiometry on wind turbine blades. Measurement 93:157–163
20. Muñoz G, Quiterio C et al (2015) Energy environment maintenance management. In: Gfeecc 2015 Gfeecc International conference forenergy, environment and commercial civilization. Chengdu, China

21. Papaelias M, Cheng L et al (2016) Inspection and structural health monitoring techniques for concentrated solar power plants. Renew Ener 85:1178–1191
22. Pliego Marugán A, García Márquez FP, Pinar Pérez JM (2016) Optimal maintenance management of offshore wind farms. Energies 9(1):46
23. Quinn TJ, Martin JE (1985) A radiometric determination of the stefan-boltzmann constant and thermodynamic temperatures between -40 degrees c and +100 degrees c. Philos Trans R Soc B Biol Sci 316(1536):85–189
24. Ramirez IS, Muñoz CQG, Marquez FPG (2017) A condition monitoring system for blades of wind turbine maintenance management. Springer

Demand Response Mechanism of a Hybrid Energy Trading Market for Residential Consumers with Distributed Generators

Nastaran Naseri, Babak Rezaee(✉), and Shahin Kazemzadeh

Industrial Engineering Department, Faculty of Engineering,
Ferdowsi University of Mashhad, PO Box: 91775-1111, Mashhad, Iran
brezaee@um.ac.ir

Abstract. Smart grid (SG) as a complex advanced electricity system is capable of facing with the growing energy demand, where energy consumers are connected not only to the conventional grid but also to some local energy markets for bidirectional energy trading. In this context, this paper presents an optimization problem the demand response problem where a hybrid energy trading system consisting of a conventional grid, a local trading system, a number of energy consumers, having distributed generators and battery, and some generator companies. In this case, it has been considered that the local trading center (LTC) is a non-profit oriented LTC which aims at benefiting the energy consumers and energy sellers. The optimization model is formulated as a mixed integer Linear Program (MILP) with a bounded number of variables and constraints. Furthermore, the solution can be obtained in polynomial time and provides the optimal scheduling for the problem such that the costs were minimized while the total demand is satisfied.

Keywords: Hybrid energy trading market · Demand response mechanism · Local trading center · Distributed generators · Battery bank

1 Introduction

Smart grid (SG) as a complex advanced electricity system is capable of facing with the growing energy demand in a reliable sustainable and economical manner [3]. Advanced two-way communication infrastructure of smart grids and efficient demand response mechanism (DR), according to which the energy consumers and energy sellers can schedule their energy consumption and energy supply, respectively, call attention to SG for achieving better performance than the conventional grid. Exclusively, bidirectional energy trading has been provided by expanded utilization of advanced smart metering systems in future SG and deployment of distributed energy sources. Accordingly, a precise design of control mechanism for both economical optimization and energy scheduling for energy consumers and sellers has been needed.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_97

Nowadays, photovoltaic (PV) generation has been considered as the most promising source of renewable energy. Moreover, the wide exploitation of PV storage facilities integrated into the utility power grid has been provided by the development of energy storage. On the top of that, the batteries can be used in order to shift peak PV generation to be in phase with peak energy demand [8].

There have been several related researches examining the energy trading mechanism and the related DR mechanism. These researches have been categorized into two groups. The first category includes those works, mainly, investigating optimal energy scheduling for energy consumers in response to the pricing of the retail market. Particularly, Mohsenian-Rad and Leon-Garcia [5], proposed an optimal and automatic residential energy consumption scheduling framework to achieve the desired trade-off between minimizing the electricity payment and minimizing the waiting time for the operation of each appliance in a household with considering a real-time pricing tariff combined with inclining block rates. They focused on a scenario which real time pricing is combined with inclining block rates to get a more balanced residential load with a low peak to average ratio. Chen et al. [2], formulated a problem to minimize the cost of energy drawn from the external grid while usage of appliances was subjected to individual delay constraints and a long-term average delay constraint. They supposed that the end user has an energy storage battery as well as an energy harvesting device lest harvested renewable energy can be stored and later used when the price is high. Kim et al. [4], formulated the energy scheduling problem as a non-cooperative game among self-interested customers, where each customer determines its own load scheduling and energy trading to maximize its own profit. They proposed a tiered billing scheme that can control the electricity rates for customers according to their different energy consumption levels in order to resolve the unfairness between heavy and light customers in the non-collaborative approach. Vaziri et al. [9], developed bi-objective integer linear programming model in order to minimize the energy costs with considering the revenue, renewable energy subsidies, and overtime costs, and minimize the displeasure of surgeons and patients in a hospital as a non-isolated micro grid. Ren et al. [7], proposed an optimization model for smart homes. The model finds out the optimal running schedule with considering annual running cost or annual CO2 emissions as the objective function to be minimized. Besides the energy flows among the equipment within the hybrid energy system, the economic information including electricity tariff and natural gas price, as well as some policy issues (e.g., buy-back price) are also accounted.

The second category subtends researches which investigated the pricing strategies for energy retailers as well as optimizing certain objectives, based on price-dependent energy scheduling. Qian et al. [6], proposed a real-time pricing scheme in order to reduce the peak-to-average load ratio through demand response management. The proposed scheme solved a two-stage optimization problem. Then, they proposed an algorithm based on simulated annealing to solve the non-convex price optimization problem. Chai et al. [1] studied DR mechanism with considering multiple utility companies and multiple residential users and a two-level game was proposed to model the interaction between these two levels. The competition among the utility companies was formulated as a non-cooperative game, while the interaction among the residential users was formulated as an evolutionary game. Then, they proved that the proposed strategies were able to make both games converge to their own equilibrium. Wu et al. [10] focused on a hybrid energy trading market consisting of an external utility company and a local trading market managed by a local trading center. First, they quantified the respective benefits of the energy consumers and the sellers from the local trading and then investigated how they can optimize their benefits by controlling their energy scheduling in response to the LTC's pricing.

The main contributions of this paper are summarized as follows. This paper first models a hybrid trading market that is comprised of multiple generator companies, multiple energy consumers, external power grid, and a local trading center. In the consumers' side, some distributed generators are considered. The consumers' appliances are categorized into three groups according to their features. The battery as a storage facility is considered for consumers so as to reduce the average peak load ratio.

The paper is organized as follows. Sect. 2 proposes the overview of the system and presents the mathematical model. Sect. 3 shows the numerical example and the results. Finally, the paper is concluded in Sect. 4.

2 Problem Definition

In this paper, a real-time scheme has been considered for a hybrid trading market in order to reduce the peak to average load ratio and maximize each user's objectives. A hybrid trading market consists of the number of Energy Consumers, Energy Sellers or Generator Companies (GenCo), a Local Trading Center (LTC) and Conventional Power Grid. Figure 1 shows a simplified illustration of the hybrid energy system.

Every energy consumer and energy seller is connected to both local trading center and the conventional power grid. Energy consumers are residential consumers which have the different type of electrical appliances, these appliances

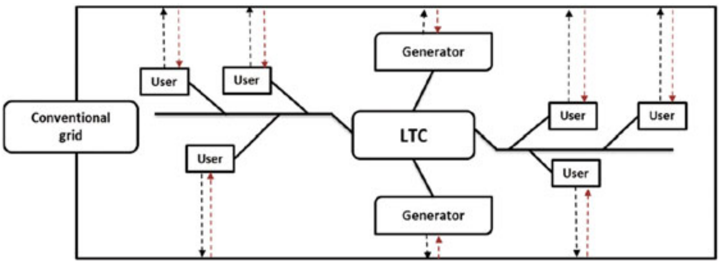


Fig. 1. A simplified illustration of the hybrid energy system

Fig. 2. An example of appliances type B

Fig. 3. An example of appliances type C

can be categorized into three main groups. The first category, A, named as non-flexible appliances, includes background appliances which are needed to be used at specific times and cannot be shifted; refrigerator, lightening appliances are some examples of this group.

The second category, B, named as semi-flexible appliances, consists of appliances which the time of using of these appliances can be shifted in a specific period of time. Figure 2 shows an example of these type of appliances. α and β are the start and end of the preferred time interval, respectively, for an appliance type B.

The third category of appliances, C, which is called as flexible appliances includes those appliances which their time of use can be shifted during a specific period of time as well as be disarticulated in the definite number of sub-activities. Figure 3 shows an example of appliance type C.

It is assumed that a number of energy consumers have Distributed Generators (DG). These DGs consist of a Photovoltaic system with a definite capacity and load availability, and a battery with a specific capacity. The battery is used to store energy surpluses during off- peak time, therefore, it can be charged via PV, LTC, and the grid; and it would be discharged only to supply power to load during peak demand hours.

In the GenCos' side, it is assumed that in every time slot h, there is a certain and constant amount of load available.

The objective of this problem is to minimize the cost of energy for consumers and maximize the income for GenCos. To implement the model, a day is divided into 24 time slots. To model the problem, the following notations are defined:

Indices:

i : Number of energy consumers $i = 1, \cdots, I$;
j : Number of generator companies $j = 1, \cdots, J$;
k : Number of sub-activities $k = 1, \cdots, K_{a_C}$;
h : Number of time slots $h = 1, \cdots, H$;
a_A : Appliances of type A $a_A = 1, \cdots, m$;
a_B : Appliances of type B $a_B = 1, \cdots, n$;
a_C : Appliances of type C $a_C = 1, \cdots, r$.

Parameters:

$[\alpha_{a_B}, \beta_{a_B}]$	: Start and finish time of appliance a_B;
$[\lambda_{a_C,k}, \gamma_{a_C,k}]$	: Start and finish time of subactivity k for appliance a_C;
l_{a_B}	: Activity time of appliance a_B;
$l'_{a_C,k}$	: Activity time of sub-activity k for appliance a_C;
$TL_{a_C,k}$	: Upper limit of the gap between two sub-activity of appliance a_C;
$TL'_{a_C,k}$	: Lower limit of the gap between two sub-activity of appliance a_C;
$d_{i,h}$	: Total energy demand of consumer i at time h;
$PV_{i,h}$	: Available load from PV system of customer i at time h;
E_{a_A}	: Total energy needed for appliances type A;
ε	: Self-discharge rate of the battery;
$\eta^{\text{discharge}}$	: Discharging efficiency of the battery;
η^{charge}	: Charging efficiency of the battery;
E^B_{capacity}	: Rated capacity of the battery (kWH);
μ	: Minimum residue coefficient of the battery
$B_{a_B,h}$	: Energy consumption of appliance a_B at time h;
M	: A large number;
$d'_{j,h}$	: Available load from GenCo j at time h;
$C_{a_C,k,h}$	: Activity k for appliance a_C at time h;
p_h	: Energy sell-out price for the grid at time h ($ /kW);
$W_{i,h}(.)$, $W_{j,h}(.)$	: A linear function for energy transmission loss ;
q_h	: Buy-back price for the grid at time h ($ /kW);
p'_h	: Energy sell-out price for LTC at time h. ($ /kW);
q'_h	: Buy-back price for LTC at time h ($ /kW)
u_h	: Buy-back price for the grid from consumers at time h ($ /kW);
u'_h	: Buy-back price for LTC from consumers at time h ($ /kW)

Decision variables:

$$ST_{a_B,h} = \begin{cases} 1, \text{ if appliance } a_B \text{ starts at time } h, \\ 0, \text{ otherwise;} \end{cases}$$

$$ET_{a_B,h} = \begin{cases} 1, \text{ if appliance } a_B \text{ finishes at time } h \\ 0, \text{ otherwise;} \end{cases}$$

$$ST'_{a_C,k,h} = \begin{cases} 1, \text{ if sub-activity } k \text{ of appliance } a_C \text{ starts at time } h, \\ 0, \text{ otherwise;} \end{cases}$$

$$ET'_{a_C,k,h} = \begin{cases} 1, \text{ if sub-activity } k \text{ of appliance } a_C \text{ finishes at time } h, \\ 0, \text{ otherwise;} \end{cases}$$

$$S_{a_B,h} = \begin{cases} 1, \text{ if the status of appliance } a_B \text{ at time } h \text{ is on} \\ 0, \text{ otherwise;} \end{cases}$$

$$S'_{a_C,k,h} = \begin{cases} 1, \text{ if the status of sub-activity } k \text{ of appliance } a_C \text{ at time } h \text{ is on} \\ 0, \text{ otherwise;} \end{cases}$$

$E_{i,h}^{\text{grid}}$: Total power purchased from the grid for customer i at time h (kW);
$E_{i,h,\text{sto}}^{\text{grid}}$: Power purchased from the grid for battery storage of customer i at time h (kW);
$E_{i,h,\text{self}}^{\text{grid}}$: Power purchased from the grid for self-use of customer i at time h (kW);
$E_{i,h}^{\text{LTC}}$: Total power bought from LTC for customer i at time h (kW);
$E_{i,h,\text{self}}^{\text{LTC}}$: Power purchased from LTC for self-use of customer i at time h (kW);
$E_{i,h,\text{sto}}^{\text{LTC}}$: Power purchased from LTC for battery storage of customer i at time h (kW);
$E_{i,h}^{PV}$: Total power of PV unit of customer i at time h (kW);
$E_{i,h,\text{self}}^{PV}$: Power out of PV unit of customer i at time h for self-use (kW);
$E_{i,h,\text{sto}}^{PV}$: Power out of PV unit of customer i at time h for battery charging (kW);
$E_{i,h,\text{salegrid}}^{PV}$: Surplus electricity from PV unit of customer i at time h sell to the grid (kW);
$E_{i,h,\text{saleLTC}}^{PV}$: Surplus electricity from PV of customer i at time h sell to LTC (kW);
$E_{i,h}^{B}$: Available power stored in the battery of customer i at time h (kW);
$E_{i,h,\text{self}}^{B}$: Power discharged from the battery of customer i at time h for self-use (kW);
$E_{i,h,\text{outLTC}}^{B}$: Power discharged from the battery of customer i at time h sell to LTC (kW);
$E_{i,h,\text{outgrid}}^{B}$: Power discharged from the battery of customer i at time h to the grid (kW);
${f_{i,h}}^{\text{grid}}$: Binary variable indicating the energy purchase from the grid for customer i
${f_{i,h}}^{B}$: Binary variable indicating the charging state of the battery of customer i
${f_{i,h}}^{\text{LTC}}$: Binary variable indicating the energy purchase from the LTC for customer i
${f_{i,h}}^{\text{grid,sto}}$: storage for battery from the grid for customer i
${f_{i,h}}^{\text{LTC,sto}}$: Binary variable indicating the energy storage for battery from the LTC for customer i
${f_{i,h}}^{PV,\text{sto}}$: Binary variable indicating the energy storage for battery from the PV for customer i
$X_{j,h}$: Total energy sold to the grid from GenCo j at time h (kW);
$Y_{j,h}$: Total energy sold to LTC from GenCo j at time h (kW)

The mathematical model for the aforementioned problem is proposed in the coming 24 h as follows. The objective function is minimizing the cost of buying Energy for customers and maximizing the income for energy consumers and sellers, respectively. These two objectives have been formulated in the following expression.

$$
\begin{aligned}
\min\ \ Z = & \sum_h \sum_i E_{i,h}^{\text{grid}} \times p_h + \sum_h \sum_i (E_{i,h}^{LTC} - W_{i,h}(E_{i,h}^{\text{LTC}})) \times {p'}_h \\
& - \sum_h \sum_i E_{i,h,\text{salegrid}}^{PV} \times u_h + \sum_h \sum_i E_{i,h,\text{saleLTC}}^{PV} \times {u'}_h \\
& - \sum_h \sum_j X_{j,h} \times q_h + \sum_h \sum_j (Y_{j,h} + W_{j,h}(Y_{j,h})) \times {q'}_h
\end{aligned} \tag{1}
$$

$$
\text{s.t.} \quad \sum_{h \geq \alpha_{a_B}}^{\beta_{a_B} - l_{a_B}} ST_{a_B,h} = 1; \forall a_B, \forall h \in H \tag{2}
$$

$$
\sum_{h \geq \alpha_{a_B} + l_{a_B}}^{\beta_{a_B}} ET_{a_B,h} = 1; \forall a_B, \forall h \in H \tag{3}
$$

$$
\sum_{h \geq \lambda_{a_C,k}}^{\gamma_{a_C,k} - {l'}_{a_C,k}} {ST'}_{a_C,k,h} = 1; \forall a_C, k, \forall h \in H \tag{4}
$$

$$
\sum_{h \geq \lambda_{a_C,k} + {l'}_{a_C,k}}^{\gamma_{a_C,k}} {ET'}_{a_C,k,h} = 1; \forall a_C, k, \forall h \in H \tag{5}
$$

$$
ST_{a_B,h} = ET_{a_B,h+l_{a_B}}; \forall a_B, \forall h \in H; \alpha_{a_B} \leq h \leq \beta_{a_B} - l_{a_B} \tag{6}
$$

$$
ST'_{a_C,k,h} = ET'_{a_C,k,h+{l'}_{a_C,k}}; \forall a_C, k, \forall h \in H; \lambda_{a_C,k} \leq h \leq \gamma_{a_C,k} - l'_{a_C,k} \tag{7}
$$

$$
S_{a_B,h} = S_{a_B,h-1} + ST_{a_B,h} - ET_{a_B,h}; \forall a_B, \forall h \in H \tag{8}
$$

$$
S'_{a_C,k,h} = S'_{a_C,k,h-1} + ST'_{a_C,k,h} - ET'_{a_C,k,h}; \forall a_C, k, \forall h \in H \tag{9}
$$

$$
\begin{aligned}
& TL'_{a_C,k} \leq h \times ST'_{a_C,k,h} - \sum_{h'=1}^{h} h' \times {ET'}_{a_C,k-1,h'} \leq TL_{a_C,k}; \\
& \forall a_C, k, k \geq 2, \forall h \in H
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
& E_{i,h,\text{self}}^{PV} + E_{i,h,\text{salegrid}}^{PV} + E_{i,h,\text{saleLTC}}^{PV} + E_{i,h,\text{sto}}^{PV} = PV_{i,h}; \\
& \forall i \in \Omega_C, \forall h \in H
\end{aligned} \tag{11}
$$

$$
\begin{aligned}
& (1 - \varepsilon) \times E_{i,h}^{B} + (E_{i,h,\text{sto}}^{\text{grid}} + E_{i,h,\text{sto}}^{\text{LTC}} + E_{i,h,\text{sto}}^{PV}) \times \eta^{\text{charge}} \\
& - \frac{(E_{i,h,\text{outgrid}}^{B} + E_{i,h,\text{outLTC}}^{B} + E_{i,h,\text{self}}^{B})}{\eta^{\text{discharge}}} = E_{i,h+1}^{B}, \forall i \in \Omega_C, \forall h \in H
\end{aligned} \tag{12}
$$

$$
\mu \times E_{\text{capacity}}^{B} \leq E_{i,h}^{B} \leq E_{\text{capacity}}^{B}; \forall i \in \Omega_C, \forall h \in H \tag{13}
$$

$$
E_{i,h,\text{sto}}^{\text{grid}} + E_{i,h,\text{sto}}^{LTC} + E_{i,h,\text{sto}}^{PV} \leq M \times f_i^{B}; \forall i \in \Omega_C, \forall h \in H \tag{14}
$$

$$
E_{i,h,\text{outgrid}}^{B} + E_{i,h,\text{outLTC}}^{B} + E_{i,h,\text{sto}}^{PV} \leq M \times (1 - {f_i}^{B}); \forall i \in \Omega_C, \forall h \in H \tag{15}
$$

$$
E_{i,h,\text{sto}}^{\text{grid}} \leq M \times f_i^{\text{grid,sto}}; \forall i \in \Omega_C, \forall h \in H \tag{16}
$$

$$E^{\mathrm{LTC}}_{i,h,\mathrm{sto}} \leq M \times f_i^{\mathrm{LTC,sto}}; \forall i \in \Omega_C, \forall h \in H \tag{17}$$

$$E^{PV}_{i,h,\mathrm{sto}} \leq M \times f_i^{PV,\mathrm{sto}}; \forall i \in \Omega_C, \forall h \in H \tag{18}$$

$$f_i^{\mathrm{grid,sto}} + f_i^{\mathrm{LTC,sto}} + f_i^{PV,\mathrm{sto}} = 1; \forall i \in \Omega_C \tag{19}$$

$$E^{\mathrm{grid}}_{i,h,\mathrm{self}} + E^{grid}_{i,h,\mathrm{sto}} \leq M \times {f_i}^{\mathrm{grid}}; \forall i \in \Omega_C, \forall h \in H \tag{20}$$

$$E^{PV}_{i,h,\mathrm{salegrid}} + E^{B}_{i,h,\mathrm{outgrid}} \leq M \times (1 - {f_i}^{\mathrm{grid}}); \forall i \in \Omega_C, \forall h \in H \tag{21}$$

$$E^{\mathrm{LTC}}_{i,h,\mathrm{self}} + E^{\mathrm{LTC}}_{i,h,\mathrm{sto}} \leq M \times {f_i}^{\mathrm{LTC}}; \forall i \in \Omega_C, \forall h \in H \tag{22}$$

$$E^{PV}_{i,h,\mathrm{saleLTC}} + E^{B}_{i,h,\mathrm{outLTC}} \leq M \times (1 - {f_i}^{\mathrm{LTC}}); \forall i \in \Omega_C, \forall h \in H \tag{23}$$

$$\begin{aligned}&\sum_{a_C}\sum_{k} C_{a_C,k,h} \times {S'}_{a_C,k,h} + \sum_{a_B} B_{a_B,h} \times S_{a_B,h} + E_{a_A} + E^{\mathrm{grid}}_{i,h,\mathrm{sto}} + E^{\mathrm{LTC}}_{i,h,\mathrm{sto}} \\ &+ E^{PV}_{i,h,\mathrm{sto}} = d_{i,h}, \forall i \in \Omega_C, \forall h \in H\end{aligned} \tag{24}$$

$$\begin{aligned}&E^{\mathrm{grid}}_{i,h,\mathrm{self}} + E^{PV}_{i,h,\mathrm{self}} + E^{B}_{i,h,\mathrm{self}} + (E^{\mathrm{LTC}}_{i,h,\mathrm{self}} - W_{i,h}(E^{\mathrm{LTC}}_{i,h,\mathrm{self}})) = d_{i,h}; \\ &\forall i \in \Omega_C, \forall h \in H\end{aligned} \tag{25}$$

$$X_{j,h} + (Y_{j,h} + W_{j,h}(Y_{j,h})) \leq d'_{j,h}; \forall j \in \Omega_G, \forall h \in H \tag{26}$$

$$\begin{aligned}&ST_{a_B,h}, ET_{a_B,h}, {ST'}_{a_C,k,h}, {ET'}_{a_C,k,h}, S_{a_B,h}, {S'}_{a_C,k,h}, f^{B}_{i,h}, f^{\mathrm{grid}}_{i,h}, f^{\mathrm{LTC}}_{i,h}, {f_{i,h}}^{PV,\mathrm{sto}}, \\ &{f_{i,h}}^{\mathrm{grid,sto}}, {f_{i,h}}^{\mathrm{LTC,sto}} \in \{0, 1\}\end{aligned} \tag{27}$$

$$\begin{aligned}&E^{\mathrm{grid}}_{i,h}, E^{\mathrm{grid}}_{i,h,\mathrm{sto}}, E^{\mathrm{grid}}_{i,h,\mathrm{self}}, E^{\mathrm{LTC}}_{i,h}, E^{\mathrm{LTC}}_{i,h,\mathrm{sto}}, E^{\mathrm{LTC}}_{i,h,\mathrm{self}}, E^{PV}_{i,h}, E^{PV}_{i,h,\mathrm{sto}}, E^{PV}_{i,h,\mathrm{self}}, E^{PV}_{i,h,\mathrm{salegrid}}, \\ &E^{PV}_{i,h,\mathrm{saleLTC}}, E^{B}_{i,h}, E^{B}_{i,h,\mathrm{self}}, E^{B}_{i,h,\mathrm{outgrid}}, E^{B}_{i,h,\mathrm{outLTC}}, X_{j,h}, Y_{j,h} \geq 0.\end{aligned} \tag{28}$$

Constraints (2)–(5) represent the start and finish time of each appliance. Continuity of appliances' activities have been avouched by constraints (6) and (7). The situation of appliances' activities (on or off) have been showed in Eqs. (8) and (9). Constraint (10) represents the limitation for the gap between each sub-activity for appliances type C; it means that the finish time of sub-activity $k-1$ and start time of sub-activity k must be in a definite limit. Constraint (11) shows that the total load available from PV system is consumed for self-use, storing in the battery, selling to the grid and LTC. Constraint (12) represents available load stored in the battery at time $h+1$ during charge and discharge process. It remarks that the stored load at time $h+1$ is equal to initial load at time h with considering the self-discharge percentage, plus incoming load from PV system, the grid, and LTC deducting the amount of load needed for the customers' demand. As the full discharge of the battery reduces its lifetime constraint (13) confines the state of charge and discharge of the battery. Constraints (14) and (15) prevent the battery from charging and discharging simultaneously. Constraints (16)–(19) prevent the process of charging from three sources (grid, LTC, and PV) simultaneously for the battery. Constraints (20)–(23) illustrate the limit of simultaneous buying and selling electricity load process for both grid and LTC. Equation (24) confines that the total energy demand for an energy consumer is coming from the total energy needed for appliances type A, B, and C, and the amount of energy needed for charging the battery. Equation (25) expresses that the energy demand can be met through the energy purchased from the grid or LTC, or the energy supplied through PV or battery; taking into consideration the amount of energy transmission loss from the LTC. The last expression (26) represents the constraint for the generators' side. Total load available from every GenCo at time can be sold to the grid or LTC. Transmission loss has been considered in the process of selling electricity to the LTC.

3 Numerical Example

In order to show the performance of the hybrid energy system consisting of the conventional grid, LTC, energy consumers with different types of appliances and GenCos, a numerical example is presented. Based on the structure of the hybrid energy system illustrated in Fig. 1, the numerical example is utilized to verify the optimization model. Figure 4 shows the energy demand for appliances of type B and the cumulative energy demand. It has been considered there are four appliances of type B.

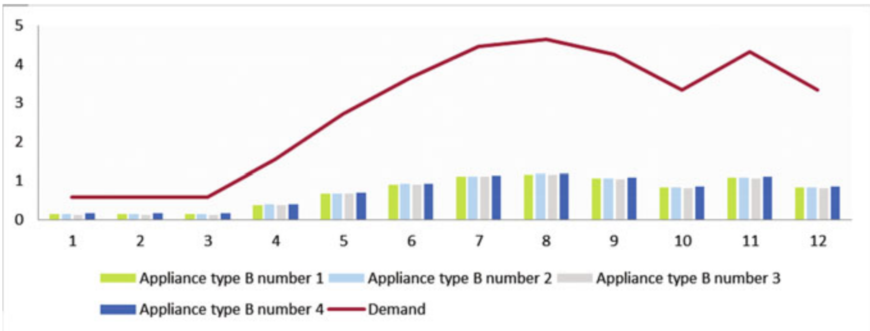


Fig. 4. Energy consumption of appliances type B and the total demand

Figure 5 shows the Energy sell-out and buy-back prices for both grid and LTC. In this example, it is assumed that the energy sell-out price for LTC is less than the energy sell-out price for the grid. Besides, the energy buy-back price for LTC is greater than energy buy-back price for the grid. The numerical example is carried out by using IBM ILOG CPLEX_Optimizer v12.3 on the PC with Intel(R) Core(TM) i7-4770 CPU@3.4 GHz and 8 GB RAM. Tables 1 and 2 represent the results. The consumer uses the PV energy and provides more energy with buying energy from LTC in order to meet energy demand in peak time.

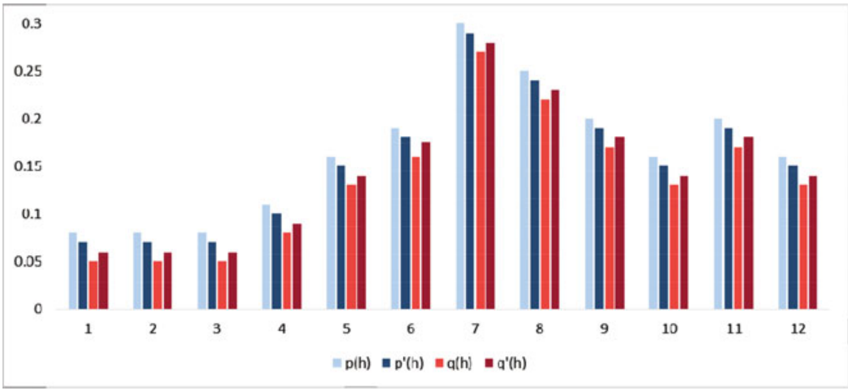


Fig. 5. Energy sell-out and buy-back prices for grid and LTC

Table 1. Energy buy-back schedule for the consumer

	Time slots											
	1	2	3	4	5	6	7	8	9	10	11	12
$E_{i,h,\text{self}}^{\text{LTC}}$	0.58	0.58	0.58	0.55	0.00	0.00	0.00	0.25	2.11	3.34	4.34	3.34
$E_{i,h,\text{self}}^{PV}$	0.00	0.00	0.00	0.99	2.72	3.66	4.46	4.41	2.16	0.00	0.00	0.00
sum	0.58	0.58	0.58	1.54	2.72	3.66	4.46	4.66	4.26	3.34	4.34	3.34

Table 2. Energy buy-back schedule for the GenCo

	Time slots											
	1	2	3	4	5	6	7	8	9	10	11	12
$Y_{j,h}$	0.13	0.13	0.13	0.51	0.82	0.99	0.79	0.95	0.88	0.54	0.99	0/89

4 Conclusion

Based on MILP theory, this paper formulated an optimization model to investigate the optimal operation strategy of the PV, battery, conventional grid, and LTC based residential hybrid energy trading system. Besides satisfying the residential electricity demands, total costs were minimized; moreover, the optimization scheme maximized the profit of GenCos. The LTC provides new opportunities for the energy consumers and GenCos to perform the local energy trading in a cooperative manner, as a result, they all can benefit. In this case, it has been considered that the LTC is a non-profit oriented LTC which aims at benefiting the energy consumers and energy sellers.

Numerical results are presented in order to validate the benefits of the considered hybrid energy market and the related DR mechanism.

References

1. Chai B, Chen J et al (2014) Demand response management with multiple utility companies: a two-level game approach. IEEE Trans Smart Grid 5(2):722–731
2. Chen S, Shroff NB, Sinha P (2013) Heterogeneous delay tolerant task scheduling and energy management in the smart grid with renewable energy. IEEE J Sel Areas Commun 31(7):1258–1267
3. Keshav S, Rosenberg C (2010) How internet concepts and technologies can help green and smarten the electrical grid. In: ACM SIGCOMM workshop on green NETWORKING 2010. New Delhi, India, pp 35–40
4. Kim BG, Ren S et al (2013) Bidirectional energy trading and residential load scheduling with electric vehicles in the smart grid. Sel Areas Commun IEEE J 31(7):1219–1234
5. Mohsenian-Rad AH, Leon-Garcia A (2010) Optimal residential load control with price prediction in real-time electricity pricing environments. IEEE Trans Smart Grid 1(2):120–133
6. Qian LP, Zhang YJA et al (2013) Demand response management via real-time electricity price control in smart grids. IEEE J Sel Areas Commun 31(7):1268–1280
7. Ren H, Wu Q et al (2016) Optimal operation of a grid-connected hybrid pv/fuel cell/battery energy system for residential applications. Energy 113:702–712
8. Torres D, Crichigno J et al (2014) Scheduling coupled photovoltaic, battery and conventional energy sources to maximize profit using linear programming. Renew Energ 72(4):284–290
9. Vaziri SM, Rezaee B, Monirian MA (2017) Bi-Objective integer programming of hospitals under dynamic electricity price. Springer, Singapore
10. Wu Y, Tan X et al (2015) Optimal pricing and energy scheduling for hybrid energy trading market in future smart grid. IEEE Trans Ind Inf 11(6):1585–1596

A Fuzzy Multi-Criteria Evaluation Method of Water Resource Security Based on Pressure-Status-Response Structure

Talat Qadeer and Zongmin Li(✉)

Business School, Sichuan University, Chengdu 610064,
People's Republic of China
lizongmin@scu.edu.cn

Abstract. In many circumstances, water resources advancement and usage has come to or surpassed the global warning level. With expanding water request and shortage, utilities require compelling techniques for ideal utilization of accessible water resources. This paper proposes a fuzzy multi-criteria evaluation method for supporting the evaluation of water resources security. Firstly, this paper establishes a novel water resource security evaluation index system based on the Pressure-Status-Response (PSR) structure. Then, a revised TOPSIS method was utilized to handle fuzzy ratings and integrate the ranking outcomes. After that, data for 9 provinces from Yellow River basin are collected and the proposed methodology is applied to test the feasibility.

Keywords: Multi-criteria decision making · Water resource security · Fuzzy evaluation · Pressure-status-response · TOPSIS

1 Introduction

Water is a fundamental resource and indispensable material basis for energy, production, human society, and economic development [8,13,19]. The pivotal role of water is beyond doubt. However, when enjoying the benefits of water, the water conflicts between human and ecosystems are becoming increasingly acute. Despite having suffered the same dilemma many times throughout history and the rapidly developed technology, sharing water resources is not something societies tend to do well [13]. The long existing water conflicts, rapid population increases, climate change, water shortage and pollution have made water resource security an indispensable component of sustainable water resource management [20]. People in developing countries are particularly at risk because of severe water pollution, environmental damages, poor water supply conditions and even social conflicts caused by water problems.

Water resource security is a concept that was proposed in the late 20th century [6,7,21]. Nowadays, water security issue is becoming more and more

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_98

prominent and has attracted a worldwide attention and emphasis [3]. The evaluation and insurance of water security are the core issues of sustainable water resources management. There are increasing studies about water resource security evaluation. For example Jiang [6] studied on water resource safety strategy for China in the 21st century. Bitterman et al. [2] proposed a conceptual framework and candidate indicators for water security and rainwater harvesting. Chen [4] studied on water resources security concept and its discussion. Xia and Zhang [21] worked on water security in north China and countermeasure to climate change and human activity. Hall and Borgomeo [9] worked on risk-based principles for defining and managing water security. Norman et al. [14] worked on water security assessment: Integrating governance and freshwater indicators. Qian and Xia [5] worked on risk assessment of water security in Haihe River Basin during drought periods based on D-S evidence theory. However, a comprehensive evaluation of water resources security is such a complex, vague and multi-level evaluation process. Multi-criteria evaluation method of water security is still worthy depth research. In addition, due to the increasingly severe climate change, water scarcity and pollution, the problem of water resources security has becoming more and more important. Therefore, this paper presents a fuzzy multi-criteria evaluation method, which include indicator system under Pressure-Status-Response framework, uncertainty rating analysis, and aggregation method of TOPSIS with fuzzy judgments.

2 Evaluation Index System Establishment

Water resource security evaluation is multi-criteria in nature. This paper establishes the evaluation index system for water resources security evaluation based on the Pressure-Status-Response (PSR) approach, which is a causal one that covers causes and effects influencing a measurable state. In this sense, three categories of indicators are distinguished [15]:

- Indicators of environmental *Pressure* describe pressures on the environment originating from human activities, including quality and quantity of natural resources (e.g., emissions, mining of raw materials, fertilizer input).
- Indicators of environmental *State* are designed to describe the status quo of the environment and the quality and quantity of resources and their changes over time (e.g., forest area, protected areas).
- Indicators of societal *Response* show to which degree society is responding to environmental changes and concerns. This could be the number and kind of measures taken, the efforts of implementing or the effectiveness of those measures. Responses may range from public (e.g., legislation, taxation, promotion) to private sector activities (e.g., reduced consumption, recycling) [12].

The originally PSR approach is in this manner changed by presenting the monetary, biological and social perspectives of water resources security into the PSR model created on the established of key variables. The model still

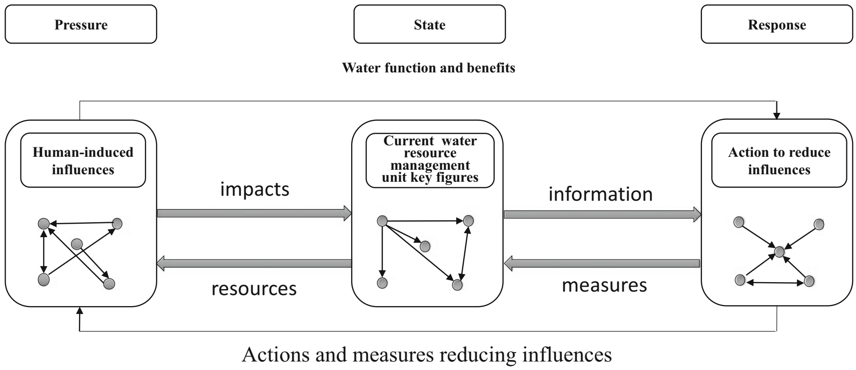


Fig. 1. Adapted pressure, state and response framework for evaluating security on the water resource management unit level

expands on the outstanding PSR structure (Fig. 1) including three boxes for Pressure, State and Response objects with the cases practically connected. Human-initiated Pressures incite impacts on State limitations by making utilitarian and monetary advantages additionally influencing the (societal) Response factors. The other way around Pressures utilize and deplete resources both subjectively and quantitatively (State) and undertaking information, choices and activities from Responses to control them. State and Response are connected by figures and material interchange on the one side and directly state-changing procedures on the other. Inside the three boxes, the structure is changed. Associations and impacts between variables are involved into evaluation by making system subsystems inside every case (Fig. 1) each pointer is appointed to each of the cases however conveying a different idea for the evaluation. Really, the examples of connotation will change among the groups since the pointers cover dissimilar PSR viewpoints for every indicator. This adjustment has been picked because a strict arrangement of indicator into boxes would mean loss of figures about the interconnections amongst indicator and their pertinence as per pressure, states and responses [18]. The interconnections between pointers are indicated by arrows in (Fig. 1).

Based on the PSR structure, we consider five criteria for water resource security, which are: water resources pressure (C_1), social-economic pressure (C_2), water resources state (C_3), social-economic state (C_4) and social-economic response (C_5).

3 Fuzzy Statement

The subjective evaluation part of water resource security evaluation, inevitably contains the evaluators' subjective judgments and preferences. This situation has further amplified the uncertainty of assessments inherent in the water resources security evaluation process. Evaluators cannot apply precise numbers to describe

Table 1. Linguistic variables and triangular IFNs for rating under the subjective evaluation criteria

Linguistic variables	Triangular IFNs
Very good (VG)	(0.9, 1.0, 1.0)
Good (G)	(0.7, 0.9, 1.0)
Medium good (MG)	(0.5, 0.7, 0.9)
Fair (F)	(0.3, 0.5, 0.7)
Medium poor (MP)	(0.1, 0.3, 0.5)
Poor (P)	(0.0, 0.1, 0.3)
Very poor (VP)	(0.0, 0.0, 0.1)

Table 2. Linguistic variables and triangular IFNs for rating the importance

Linguistic variables	Triangular IFNs
Very high (VH)	(0.7, 0.9, 1.0)
High (H)	(0.6, 0.7, 0.8)
Medium high (MH)	(0.4, 0.5, 0.6)
Medium (M)	(0.1, 0.3, 0.5)
Medium low (ML)	(0.1, 0.2, 0.3)
Low (L)	(0.0, 0.1, 0.2)
Very low (VL)	(0.0, 0.0, 0.2)

their assessments, however, they can utilize linguistic variables according to their professional knowledge and experience. Hence, the concept of fuzzy numbers can be integrated into the multi-criteria evaluation of water resources security.

Evaluators first make their own judgments of water resource security based on subjective evaluation criteria C_1–C_5. Ratings under subjective evaluation criteria are considered as linguistic variables. A linguistic variable is a variable whose value is a natural language phrase. It is very useful in dealing with situations which are ill-defined to be described properly in conventional quantitative expressions. Water resource security performance under each subjective evaluation criteria can be expressed on a 7-point rating scale: "very good","good", "medium good", "fair", "medium poor", "poor", and "very poor". Such linguistic variables are converted into triangular intuitionistic fuzzy numbers (IFNs) [1] as shown in Table 1. The linguistic variables and triangular IFNs for rating the importance are shown in Table 2. Figures for membership functions are shown in Figs. 2 and 3. IFNs are commonly used for solving decision-making problems, where the available information is imprecise. There are different shapes or forms of IFNs, among those, trapezoidal IFNs and triangular IFNs are the most commonly used. For example, Shaw and Roy [17] used trapezoidal IFNs for analysing fuzzy system reliability, while Vahdani et al. [22] applied triangular IFNs to fuzzy

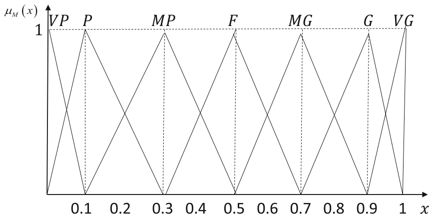


Fig. 2. Figure for membership functions of linguistic variables of rating under the subjective evaluation criteria

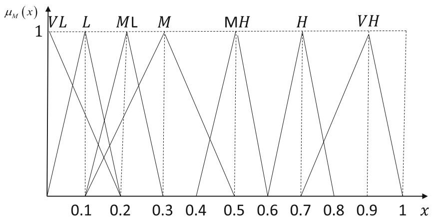


Fig. 3. Figure for membership functions of linguistic variables for rating under the importance

group decision-making problems with an application to the contractor selection. This paper chooses to use the triangular IFNs because of their conceptual and computational simplicities [11].

4 Aggregation Method

TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution) method proposed by Hwang and Yoon [10] is adopted to aggregate all ratings. The basic principle of TOPSIS is that the chosen alternative must have the shortest distance from the ideal solution and the farthest distance from the negative-ideal solution [16]. The original TOPSIS is only applicable for crisp ratings. In this paper, all ratings are represented by triangle fuzzy numbers. Therefore, this paper extents the original TOPSIS to fuzzy environment. The detailed processes are listed as follows.

Suppose the fuzzy ratings under all criteria are denoted as $\boldsymbol{A} = [a_{ij}], (i = 1, \cdots, m, j = 1, \cdots, n)$.

Step 1. Normalize the evaluation index as: Compute the normalized fuzzy decision matrix ($\boldsymbol{V} = [v_{ij}]$). The normalized value n_{ij} is calculated as:

$$v_{ij} = \frac{a_{ij}}{\sqrt{\sum a_{ij}^2}}. \tag{1}$$

Step 2. Calculate entropy measure of every index the following equation:

$$p_{ij} = \frac{a_{ij}}{\sum_{i=1}^{m} a_{ij}}, \tag{2}$$

$$E_j = -K \sum_i [P_{ij} \times \ln P_{ij}], \quad \text{where} \quad K = \frac{1}{\ln(m)}. \tag{3}$$

Step 3. Define the divergence through:

$$d_j = 1 - E_j. \tag{4}$$

Step 4. Obtained the normalized weights of indexes as:

$$W_j = \frac{d_i}{\sum d_j}. \tag{5}$$

Step 5. Let $\tilde{a} = (a_1, a_2, a_3)$ and $\tilde{b} = (b_1, b_2, b_3)$ be two triangular fuzzy numbers, then the vertex method is defined to calculate the distance between them:

$$d(\tilde{a}, \tilde{b}) = \sqrt{\frac{1}{3}[(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2]}. \tag{6}$$

Step 6. Determine the ideal and negative-ideal solution:

$$v_j^+ = v_i^+, \cdots, v_n^+ = [(\max v_{ij} | i \in I), (\min v_{ij} | i \in I)], \tag{7}$$

$$v_j^- = v_i^-, \cdots, v_n^- = [(\min v_{ij} | i \in I), (\max v_{ij} | i \in I)]. \tag{8}$$

Step 7. Calculate the separation measures, using the n-dimensional Euclidean distance. The separation of each alternative from the ideal solution is given as:

$$d_j^+ = \sqrt{\sum_{j=i}^{n} (v_{ij} - v_j)^2}, i = (1, 2, \cdots, m), \tag{9}$$

$$d_j^- = \sqrt{\sum_{j=i}^{n} (v_{ij} - v_j)^2}, i = (1, 2, \cdots, m). \tag{10}$$

Step 8. Calculate the relative closeness to the ideal solution. The relative closeness of the alternative a_j with respect to A^* is defined as:

$$C.L_i = \frac{d_j^-}{d_i^- + d_i^+}, I = 1, 2, \cdots, m. \tag{11}$$

Step 9. Rank the preference order. By ordering $C.\ L_i$ in descending order, provinces can be ranked from the best to the worst.

5 Case Study

This study presents a case study of 9 provinces from Yellow River basin. These 9 provinces of China are Qinghai, Sichuan, Gansu, Ningxia, Inner Mongolia, Shaanxi, Shanxi, Henan and Shandong. The original data are from Liu et al. [13]. Based on the original data, we have experts to make their fuzzy judgements for all the criteria (C_1–C_5). The decision table is Table 3.

We also have experts give their ratings about the importance of all criteria as shown in Table 4.

Based on Eqs. (1) and (2), we obtained results shown in Tables 5 and 7.

By using Eq. (3) results are $E_1, E_2, E_3, E_4, E_5 = (0.9, 0.9, 0.6), (0.8, 0.9, 0.9), (0.8, 0.9, 0.9), (0.6, 0.8, 0.9), (0.8, 0.8, 0.9)$ respectively (Table 6).

Table 3. Fuzzy judgements for all criteria

Criteria alternatives	Water resources pressure (C_1)	Socio-economic pressure (C_2)	Water resources state (C_3)	Socio-economic state (C_4)	Socio-economic response (C_5)
Qinghai	MH	ML	G	MP	MG
Sichuan	MH	L	G	P	Vp
Gansu	H	H	P	P	F
Ningxia	VH	ML	MG	G	F
Inner Mongolia	H	MH	F	G	VP
Shaanxi	H	ML	MG	P	MP
Shanxi	H	L	MP	VP	F
Henan	MH	MH	MP	MP	MP
Shandong	MH	F	VP	P	P

Table 4. Decision matrix for importance

Criteria alternatives	Water resources pressure (C_1)	Socio-economic pressure (C_2)	Water resources state (C_3)	Socio-economic state (C_4)	Socio-economic response (C_5)
Qinghai	MH	ML	G	MP	MG
Sichuan	MH	L	G	P	Vp
Gansu	H	H	P	P	F
Ningxia	VH	ML	MG	G	F
Inner Mongolia	H	MH	F	G	VP
Shaanxi	H	ML	MG	P	MP
Shanxi	H	L	MP	VP	F
Henan	MH	MH	MP	MP	MP
Shandong	MH	F	VP	P	P

Table 5. The quantitative decision matrix

Criteria alternatives	C_1^-	C_2^-	C_3^+	C_4^+	C_5^-
Qinghai	0.4, 0.5, 0.6	0.1, 0.2, 0.3	0.7, 0.9, 1.0	0.5, 0.7, 0.9	0.5, 0.7, 0.9
Sichuan	0.4, 0.5, 0.6	0.0, 0.1, 0.2	0.7, 0.9, 1.0	0.0, 0.1, 0.3	0.0, 0.0, 0.1
Gansu	0.6, 0.7, 0.8	0.6, 0.7, 0.8	0.0, 0.1, 0.3	0.0, 0.1, 0.3	0.3, 0.5, 0.7
Ningxia	0.7, 0.8, 0.9	0.1, 0.2, 0.3	0.5, 0.7, 0.9	0.7, 0.9, 1.0	0.3, 0.5, 0.7
Inner Mongolia	0.6, 0.7, 0.8	0.4, 0.5, 0.6	0.3, 0.5, 0.7	0.7, 0.9, 1.0	0.0, 0.0, 0.1
Shaanxi	0.6, 0.7, 0.8	0.1, 0.2, 0.3	0.5, 0.7, 0.9	0.0, 0.1, 0.3	0.1, 0.3, 0.5
Shanxi	0.6, 0.7, 0.8	0.0, 0.1, 0.2	0.1, 0.3, 0.5	0.0, 0.0, 0.1	0.3, 0.5, 0.7
Henan	0.4, 0.5, 0.6	0.4, 0.5, 0.6	0.1, 0.3, 0.5	0.1, 0.3, 0.5	0.1, 0.3, 0.5
Shandong	0.4, 0.5, 0.6	0.1, 0.3, 0.5	0.0, 0.0, 0.1	0.0, 0.1, 0.3	0.0, 0.1, 0.3
$\sum a_{ij}$	4.7, 5.6, 6.5	1.8, 2.8, 3.8	2.9, 4.4, 5.9	2, 3.2, 4.7	1.6, 2.9, 4.5
$\sqrt{\sum_{i=1}^{m} a_{ij}^2}$	1.6, 1.9, 2.2	0.8, 1.1, 1.4	1.3, 1.7, 2.2	1.1, 1.5, 1.9	0.7, 1.2, 1.7

Note: − and + signs denote the negative and positive criteria respectively.

Table 6. Normalize the evaluation index

Criteria alternatives	C_1^-	C_2^-	C_3^+	C_4^+	C_5^-
Qinghai	0.3, 0.3, 0.3	0.1, 0.2, 0.2	0.6, 0.5, 0.5	0.5, 0.5, 0.5	0.7, 0.6, 0.0
Sichuan	0.3, 0.3, 0.3	0.0, 0.1, 0.1	0.6, 0.5, 0.5	0.0, 0.1, 0.2	0.0, 0.0, 0.1
Gansu	0.4, 0.4, 0.4	0.7, 0.6, 0.6	0.0, 0.1, 0.1	0.0, 0.1, 0.2	0.4, 0.4, 0.4
Ningxia	0.4, 0.4, 0.4	0.1, 0.2, 0.2	0.4, 0.4, 0.4	0.6, 0.5, 0.5	0.4, 0.4, 0.4
Inner Mongolia	0.4, 0.4, 0.4	0.5, 0.5, 0.4	0.2, 0.3, 0.3	0.6, 0.5, 0.5	0.0, 0.0, 0.1
Shaanxi	0.4, 0.4, 0.4	0.1, 0.2, 0.2	0.4, 0.4, 0.4	0.0, 0.1, 0.2	0.1, 0.3, 0.3
Shanxi	0.4, 0.4, 0.4	0.0, 0.1, 0.1	0.1, 0.2, 0.2	0.0, 0.0, 0.1	0.4, 0.4, 0.4
Henan	0.3, 0.3, 0.3	0.5, 0.5, 0.4	0.1, 0.2, 0.2	0.9, 0.2, 0.3	0.1, 0.3, 0.3
Shandong	0.3, 0.3, 0.3	0.1, 0.3, 0.4	0.0, 0.0, 0.1	0.0, 0.1, 0.2	0.0, 0.1, 0.2

Note: − and + signs denote the negative and positive criteria respectively.

By using Eq. (4) divergence are $d_1, d_2, d_3, d_4, d_5 = (0.0, 0.0, 0.4), (0.3, 0.1, 0.1), (0.2, 0.1, 0.1), (0.4, 0.2, 0.1), (0.3, 0.2, 0.1)$ respectively.

According TOPSIS, we obtained weights: $W = [(0.005, 0.005, 0.154), (0.099, 0.038, 0.020), (0.084, 0.050, 0.024), (0.184, 0.095, 0.038), (0.104, 0.068, 0.033)]$.

By using Eq. (5) the normalized weights are $W_1, W_2, W_3, W_4, W_5 = (0.0, 0.0, 0.2)$,
$(0.1, 0.0, 0.0), (0.1, 0.1, 0.0), (0.2, 0.1, 0.0), (0.1, 0.1, 0.0)$ respectively.

Table 7. Calculate entropy measure

Criteria alternatives	C_1^-	C_2^-	C_3^+	C_4^+	C_5^-
Qinghai	0.1, 0.1, 0.1	0.1, 0.1, 0.1	0.2, 0.2, 0.2	0.3, 0.2, 0.2	0.3, 0.2, 0.2
Sichuan	0.1, 0.1, 0.1	0.0, 0.0, 0.1	0.2, 0.2, 0.2	0.0, 0.0, 0.1	0.0, 0.0, 0.0
Gansu	0.1, 0.1, 0.1	0.3, 0.3, 0.2	0.0, 0.0, 0.01	0.0, 0.0, 0.1	0.2, 0.2, 0.2
Ningxia	0.2, 0.1, 0.1	0.1, 0.1, 0.1	0.2, 0.2, 0.2	0.4, 0.3, 0.2	0.2, 0.2, 0.2
Inner Mongolia	0.1, 0.1, 0.1	0.2, 0.2, 0.2	0.1, 0.1, 0.1	0.4, 0.3, 0.2	0.0, 0.0, 0.0
Shaanxi	0.1, 0.1, 0.1	0.1, 0.1, 0.1	0.2, 0.2, 0.2	0.0, 0.0, 0.1	0.1, 0.1, 0.1
Shanxi	0.1, 0.1, 0.1	0.0, 0.0, 0.1	0.0, 0.1, 0.1	0.0, 0.0, 0.0	0.2, 0.2, 0.2
Henan	0.1, 0.1, 0.1	0.2, 0.2, 0.2	0.0, 0.1, 0.1	0.0, 0.1, 0.1	0.1, 0.1, 0.1
Shandong	0.1, 0.1, 0.1	0.1, 0.1, 0.1	0.0, 0.0, 0.0	0.0, 0.0, 0.1	0.0, 0.0, 0.1

Note: − and + signs denote the negative and positive criteria respectively.

After the weighted normalized decision matrix we obtain a fuzzy weighted decision table. Taking Qinghai as example, by applying the Eq. (6) results are given below:

$$\begin{aligned}
D_1^+ = &+\sqrt{\frac{1}{3}[(0-0.001)^2+(0-0.001)^2+(0-0.042)^2]}\\
&+\sqrt{\frac{1}{3}[(0-0.012)^2+(0-0.007)^2+(0-0.004)^2]}\\
&+\sqrt{\frac{1}{3}[(1-0.047)^2+(1-0.026)^2+(1-0.011)^2]}\\
&+\sqrt{\frac{1}{3}[(1-0.083)^2+(1-0.045)^2+(1-0.019)^2]}\\
&+\sqrt{\frac{1}{3}[(1-0.071)^2+(1-0.040)^2+(1-0.002)^2]}\\
=&2.756,\\
D_1^- = &+\sqrt{\frac{1}{3}[(1-0.001)^2+(1-0.001)^2+(1-0.042)^2]}\\
&+\sqrt{\frac{1}{3}[(1-0.012)^2+(1-0.007)^2+(1-0.004)^2]}\\
&+\sqrt{\frac{1}{3}[(0-0.047)^2+(0-0.026)^2+(0-0.011)^2]}\\
&+\sqrt{\frac{1}{3}[(0-0.083)^2+(0-0.045)^2+(0-0.019)^2]}\\
&+\sqrt{\frac{1}{3}[(0-0.071)^2+(0-0.040)^2+(0-0.002)^2]}\\
=&2.112.
\end{aligned}$$

Calculate the relative closeness to the ideal solution using Eq. (4), Qinghai's results are given below:

$$CL_{\text{Qinghai}} = \frac{2.112}{2.112 + 2.756} = 0.434.$$

By the TOPSIS method, we get the rank of 9 provinces as shown in Table 8.

Table 8. Table of ranks

Alternatives	TOPSIS index	Rank
Qinghai	0.434	1
Sichuan	0.403	5
Gansu	0.394	9
Ningxia	0.417	2
Inner Mongolia	0.405	4
Shaanxi	0.399	7
Shanxi	0.402	6
Henan	0.409	3
Shandong	0.396	8

Based on the ranking result of TOPSIS method, the water resources security degrees of the 9 provinces.

6 Conclusions

This paper presented a fuzzy multi-criteria evaluation TOPSIS method for the water resource security evaluation. 9 provinces within the Yellow River basin was ranked for the water resources security degree based on the proposed method. The structure of "pressure-state-response" was embedded in developing the methodology. A fuzzy multi-criteria was proposed not just to deal with the evaluation created on the established indicators, additionally to tackle the inherent suspicions. Based on the case study results, Qinghai, Ningxia, Henan, Inner Mongolia and Sichuan are ranked top 5 of the water security degree, with TOPSIS index greater than 0.4. Future research will be focused on the uncertainty analysis, innovative evaluation index system and aggregation methods of water resource security evaluation.

Acknowledgements. We are thankful for financial support from the Research Center for Systems Science & Enterprise Development (Grant No. Xq15C01), the National Natural Science Foundations (Grant No. 71671118, Grant No. 71601134) and Project funded by China Postdoctoral Science Foundation.

References

1. Atanassov KT (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20:87–96
2. Bitterman P, Tate E et al (2016) Water security and rainwater harvesting: a conceptual framework and candidate indicators. Appl Geogr 76:75–84
3. Cook C, Bakker K (2012) Water security: debating an emerging paradigm. Glob Environ Change 22:94–102
4. Chen SJ (2004) Water resources security concept and its discussion. China Water Res 17:13–15
5. Dong QJ, Liu X (2014) Risk assessment of water security in Haihe River Basin during drought periods based on D-S evidence theory. Water Sci 7:119–132
6. Jiang WL (2001) Study on water resource safety strategy for China in the 21st century. Adv Water Sci 1:66
7. Jia SF, Zhang JY, Zhang SF (2002) Regional water resources stress and water resources security appraisement indicators. Progr Geogr 21:538–545
8. Koehler A (2008) Water use in LCA: managing the planet's freshwater resources. Int J Life Cycle Assess 13:451–455
9. Hall J, Borgomeo E (2013) Risk-based principles for defining and managing water security. Philos Trans R Soc London A Math Phys Eng Sci 371:20120407
10. Hwang CL, Yoon K (1981) Multiple attributes decision making methods and applications. Springer, Heiddberg
11. Li Z, Liechty M et al (2014) A fuzzy multi-criteria group decision making method for individual research output evaluation with maximum consensus. Knowl Based Syst 56:253–263
12. Linser S (2001) Critical analysis of the basics for the assessment of sustainable development by indicators. Schriftenreihe Freiburger Forstliche Forschung, Bd. 17, Freiburg
13. Liu KK, Li CH et al (2014) Comprehensive evaluation of water resources security in the Yellow River basin based on a fuzzy multi-attribute decision analysis approach. Hydrol Earth Syst Sci 18:1605–1623
14. Norma ES, Dunn G et al (2013) Water security assessment: integrating governance and freshwater indicators. Water Res Manag 27:535–551
15. OECD (1993) Core set of indicators for environmental performance reviews: a synthesis report by the group on the state of the environment. Environment monographs, vol. 83. Organization for Economic Co-operation and Development, Paris
16. Shirouyehzad H, Dabestani R (2011) Evaluating projects based on safety criteria: using TOPSIS. In: 2011 2nd International conference on construction and project management IPEDR, vol. 15. Singapore
17. Van Raan AFJ (1996) Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. Scientometrics 36:397–420
18. Vacik H, Wolfslehner B et al (2007) Integrating the DPSIR-approach and the analytic network process for the assessment of forest management strategies. In: Reynolds K, Rennolls K et al (eds) Sustainable forestry: from monitoring and modelling to knowledge management and policy science. CABI Publishing
19. Voeroesmarty CJ, McIntyre PB et al (2010) Global threats to human water security and river biodiversity. Nature 467:555–561

20. Wolf A (2007) Shared waters: conflict and cooperation. Annu Rev Environ Res 32:269–279
21. Xia J, Zhang YY (2007) Water security in north China and countermeasure to climate change and human activity. Phys Chem Earth Parts A/B/C 33:359–363
22. Xu JP, Li ZM et al (2013) Multi-attribute comprehensive evaluation of individual research output based on published research papers. Knowl Based Syst 43:135–142

Management of Technological Modes of System Distributed Generation Electric Energy on the Basis of Daily Schedules of Electric Loadings

Abdulla Arifjanov[1(✉)] and Romen Zakhidov[2]

[1] Center for the Development of Software Products and Hardware-Software Complexes, Tashkent University of Information Technologies, Tashkent, Uzbekistan
arifjanov@yandex.ru

[2] Institute of Power Engineering and Automation, Academy of Sciences of Uzbekistan, Tashkent, Uzbekistan

Abstract. Currently the concept of distributed power with application of renewable energy sources is gaining popularity and is becoming not only a supplement but also a real alternative of traditional energy supply from centralized networks. The distributed generators of electricity integrated into Unified Energy System (UES), represent difficult system in which communication between observable parameters and a system condition has difficult and ambiguous character. Therefore, the approach to management in the automatic control systems, based on management on output parameters, is inadequate as a matter of fact. In this case formation of operating influence should be carried out on the basis of data about a current condition of object of management. The daily schedules of electric load allow properly assess the mode of the work of electrical equipment and the enterprise as a whole, identify the bottlenecks and reserves, to set the optimal mode of the work. In this paper are considered an issue of the development of methods for the estimation of electrical loads for the various types of electricity customers in the residential and public buildings and addresses issues of the integration of distributed generators to the UES and ensuring the optimal management of technological modes of their work.

Keywords: Electric power system · Hybrid system of power supply · Renewable energy sources · Distributed generation electric energy · Electrical loads · Daily schedule load · Management

1 Introduction

The future power supply system should combine large power stations without which are problematic electro supply of large consumers and maintenance of growth of power consumption. Large power stations allow raising voltage with

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_99

generating to voltage of the main network on which there is a transport of the electric power to the large centers of consumption. Installations of the distributed generation, including on renewable energy sources (RES), work through distributive networks. The following level of a power supply system of the future will make mini-and micro-installations (mini-and micro-hydroelectric power station, wind generators, solar energy installations, fuel elements, etc.) which are connected on a low voltage and are established at small consumers, for example, in separate houses or on small enterprises, including agricultural [1,7].

In this regard, is currently the concept of distributed power with application of renewable energy sources (RES) is gaining popularity and is becoming not only a supplement but also a real alternative of traditional energy supply from centralized networks.

Distributed generation systems represent a hybrid power system consisting of various energy sources and are built in close proximity to the customer and to the maximum extent take into account the individual characteristics of the consumer's by power and profile (Fig. 1). They have such advantages as controllability, reliability, economy (including by eliminating the losses of transporting), scalability and flexibility [7].

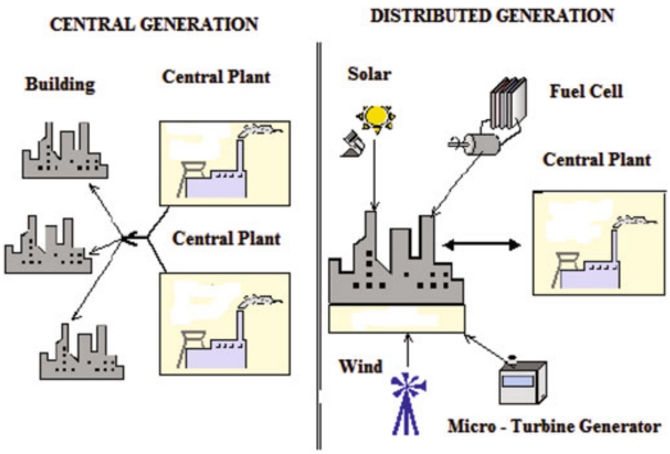


Fig. 1. A structural diagram a centralized and distributed power generation

In these systems electric energy is generated with use of solar photo-electric panels, wind turbines or other energy installations. Generating of thermal energy for systems of heating, hot water supply and technological processes is carried out with use of solar collectors (flat and vacuum tubular), geothermal systems, and also other converters of thermal energy. The combination of various renewable energy sources is not only presence of such elements, as solar collectors, photo-electric panels, wind turbines, thermal pumps, but also use of a uniform control system for maintenance of effective work of these elements that makes a basis of stable hybrid system of power supply.

As of today the distributed generation systems in most cases presented in the form of autonomous energy centers, as the connection of distributed generators to centralized networks so far is limited by the absence or imperfection of the regulatory framework and effective methods of controls of technological modes.

For an effective utilization of hybrid system of power supply with application of renewable energy sources (RES) it is necessary to co-ordinate intensity of receipt of energy with its expense consumers. Such coordination can be carried out on the basis of monitoring of the current information on receipt and power consumption and the management which is carried out by uniform system, constructed on the basis of modern information-communication technologies [2].

Control of the distributed generators is carried out as control of a micro network (micro grid) of the "virtual" power stations integrated into a global network (Unified Energy System-UES) and, thereby, in the electric power and capacity market that will promote increase of a role of the consumer in management of a power supply system [2].

On a way of creation of "virtual" power stations it is necessary to solve a number of the organizational-technological problems, one of which is the problem of connection of the distributed generators to a uniform network and management of their work. The distributed generators of electricity integrated into UES, represent difficult system in which communication between observable parameters and a system condition has difficult and ambiguous character. Therefore, the approach to management in the automatic control systems, based on management on output parameters, is inadequate as a matter of fact. In this case formation of operating influence should be carried out on the basis of data about a current condition of object of management [2].

The specific issues on usage of renewable energy sources as a distributed generator, as well the approach to the synthesis of the operational level of the active-adaptive system of automatic operation control of the combined plant as part of the UES were reviewed in the report [2].

The effectiveness of control by system of distributed generation of electrical energy, comprising renewable energy sources (RES), energy storages and intelligent control system is achieved through the ability to complex influence on the technological parameters of the generator modes.

For example, voltage regulation apart from the use of traditional means, may be implemented by change the generation of active or reactive power, by exposure on the energy storage means or by management of load.

This set of measures together with hardware and software represents the control system of system of distributed generation of electric energy.

The main purpose of this control system is the integration of distributed generators to the UES and to ensure the optimal management of technological modes of their work to meet the needs of consumers in the electrical energy.

Needs for electric power supply are characterized by the schedule of electrical loads. Dependence of load on time of day is called the daily schedule of electric load. It can be compiled for individual and group of electricity consumers, and for the Energy System, which includes a power plant or for large of the incorporated power system [6].

The daily schedules of electric load allow properly assess the mode of the work of electrical equipment and the enterprise as a whole, identify the bottlenecks and reserves, to set the optimal mode of the work.

Most fundamentally change in the electrical load associated with communal general needs. Fig. 2 represents the daily schedule, from which we see that the electrical load in the winter more than the summer and is sharply reduced at night hours [6]. The smallest of its value called of the minimum load. In the afternoon and evening hours is observed increasing of load and a more considerable change - in winter. There are two maximum loads-morning and evening. Schedule of electrical loads should be provided ("covered") on a mandatory basis. Therefore, aspire carry out all necessary renovations during the summer in order to virtually all equipment of power plant could be used to provide winter peak. This maximum is called the peak load.

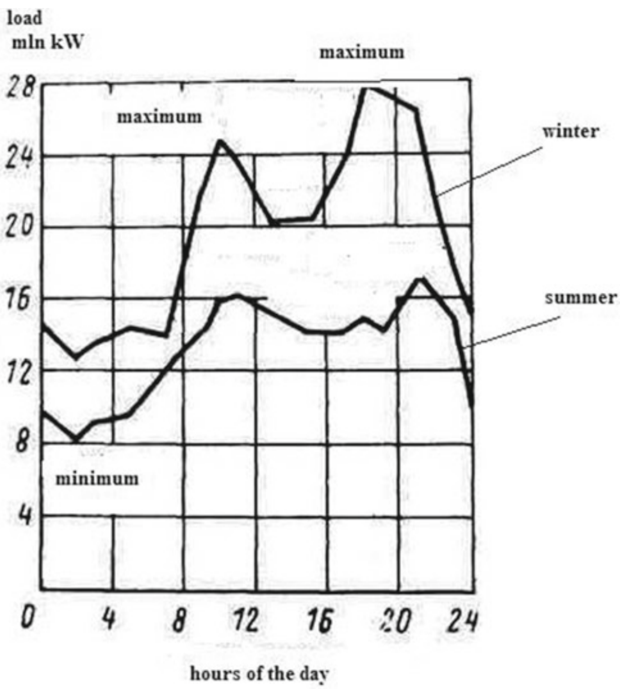


Fig. 2. Daily schedule of electric load

Calculation of the urban network load includes determining the load of individual consumers (residential buildings, public buildings, municipal services, etc.) and electrical elements of the system (distribution lines, transformer substations, distribution centers, power centers, etc.) [3].

In connection with the appearance of part of the population the possibility of use in the home of a wide range of modern household appliances and equipment, as well as in connection with the construction of urban and rural buildings on individual projects with luxury apartments, new specific standards of electrical loads should be determined based on the actual measurements electrical loads, taking into account their probabilistic nature and features of modern luxury apartments in new buildings of the city, analysis of electrical appliances market development and technology, the degree of saturation of apartments, both now and in the future.

In the cities, the problem of reliable determination of residential and public buildings loads, and provide energy savings is especially difficult, because of population growth and an increase the number of used electrical appliances. Along with the increasing number of electrical appliances used by the population grows and their installed capacity.

Modern residential buildings full of a lot of different power consumers. These include lighting and household appliances and power equipment. There is a constant process of improving housing comfort, and this in turn increases the amount of domestic power consumers and increased domestic electricity consumption. Increasing the number of storeys homes tightens the requirements for reliability and continuity of supply of electric networks.

The same features are characteristic and for public buildings, which include: a variety of institutions and the organization of management, finance, credit, state monopolies, education; pre-schools, libraries, archives, enterprises of trade, public catering, consumer services; hotels, hospitals, museums, entertainment and sports facilities.

In this regard, the reliable definition of electric loadings of residential and public buildings in the cities became very actual problem on which solution the end results of actions for energy saving in many respects depend. The normative document The electric facilities of residential and public buildings. Construction norms and rules using to calculate of loadings of residential buildings doesn't solve the designated problem, because the specific norms of electric loadings given in this document for various types of consumers of energy have been defined in the 80s last century and don't correspond to today's condition of power consumption in residential and public buildings. In many cases the calculated values of electric loadings of residential and public buildings defined according to this normative document with use of various correction coefficients aren't true [7].

2 The Problem Statement

Analysis of the existing regimes of power consumption in the urban electric network shows the need to develop new accounting standards specific to electrical loads for different types of residential and public buildings consumers, taking into account the current realities of power consumption in the residential and public sector [3–5].

Estimated electrical loads and dynamics of their growth in the future is the foundation that determines the nature of construction and the development of urban electrical networks. The electrical loads of houses and public buildings are random and depend on some of factors: the way of life of the various families, the number of power consumers and power and others. Therefore, the basis for determining the load is used the probabilistic and statistical approach to the load as a non-stationary random variable. Hence, the estimated electrical load or network element assumes to be probable maximum load value for an interval of 30 min.

To obtain reliable data while a design of standard charts and determine their numerical characteristics it is necessary correctly handle the experimental results, based on the position of mathematical statistics and probability theory. It should be borne in mind that, according to the law of large numbers and the theory of probability, the results determine the average aggregate will be valid when the number of tests or, in other words, the number of surveyed members together will be quite large. On the other hand, with an increase in the number of aggregate members, their examination is a huge work, and we have a problem of determining a sufficient number of members in the aggregate, that allow to obtain mean values with sufficient accuracy.

Therefore, for an assessment of this loading very often used some generalized indicators, coefficients, specific loadings and specific expenses of the electric power.

The existing algorithms of load calculation of the industrial power systems, determination of the maximum loads and the choice of electric equipment don't consider dynamics of growth and character of household loading, laws of its functioning, do not allow to take into account in calculating of the actual loading the relationship of load changes on a day of the week, time of day. Despite the large number of works on the subject, the daily schedule load model of residential and public buildings and their practical implementation are not well developed. There is not a program that gives to specify the load calculation data of residential and public buildings on the current state of the loads.

The method widely used for determining the maximum electrical load of electrical networks are based on the measurement of the average load of electrical consumers over a given time period t ($t = 8$ h or 0.5 h) with variable initial measuring point. The total number of electrical consumers, on which measurements are made should not be less than 20% of the total number of electrical consumers connected to the electric network (but not less than 15). Measurements should be carried out repeatedly and for a long time [2,7]. All these methods are characterized by a long time and a low accuracy of measurements; by a significant difference from the actual loads, especially the total; by significant errors in calculation (do not include the probabilistic nature of the electrical loads in the urban network). These methods also do not take into account the time factor of maximum electric load of each customer. In addition, these techniques are designed for installations with a regular operating mode and require large amounts of additional measurements and not possible to determine the desired value for a predetermined time period including at least three units.

The way of definition of the maximum load of electric consumers [6] according to which measure once individual loadings in various technological operating modes, total time of each technological mode for basic time and calculate the load group with a probability of exceeding no more than the required conditions of the problem.

However, this method does not provide sufficient accuracy while determining the estimated load for a group of different types of electric consumers, in particular, if even in the same technological modes individual electrical loads are not constant.

3 A New Method of Solving the Problem

A method for determining the maximum electrical load of electrical networks and individual electrical consumers patented by one of author of this paper [5].

This method significantly reduces the numbers and time of measurements and improve the accuracy and reliability of determining the maximum electrical load of consumers.

The essence of a way is electric loads of the certain consumers are measured in various technological operating modes for basic time. Then from set of the measured loads the maximum size of load of the certain consumers is determined. The maximum value of the electric load is defined as the highest value for all of the measured periods of typical daily schedules, with a maximum electrical load in a time interval 15-minute and 30-minute. Further calculation of loads of a network is conducted in a general view on a formula [4].

Let's us consider the essence and use of this new method. The essence of the new method for calculating the load of 0.4 kV and 6–10 kV electrical networks is to determine estimated load for each time period of the day according to the on typical daily schedule of customer load. It is found by summing of the mathematical expectations and the mean square deviations of all consumers. For the calculated maximum value of the total load is accepted its greatest value from all the calculating for hour or half-hour periods of the daily schedule.

The mathematical expectation for the construction of a typical load graph of homogeneous consumers will be equal [4]:

$$\overline{P_{cpi}} = \frac{\sum_{\beta=1}^{m} \times \sum_{\gamma=1}^{l} \times \sum_{j=1}^{n} Y_{ij\gamma\beta}}{m\,\ln},$$

where

P_{cpi} : mathematical expectation of the random load;

Y_{ij} : matrix of initial values of the loads daily schedule;

n : the number of points measured value of the daily schedule for a half-hour or hour interval;

l : the number of load measurements (in days) in each point at electrical network;

m : the number of homogeneous consumers, accepted as experimental.

So, the root-mean-square deviation equals:

$$\sigma_{\overline{p}i} = \sqrt{D_i} = \sqrt{\frac{\sum_{\beta=1}^{m} \times \sum_{\gamma=1}^{l} \times \sum_{j=1}^{n} Y_{ij\gamma\beta}^{2}}{m\,\ln} - \left(\frac{\sum_{\beta=1}^{m} \times \sum_{\gamma=1}^{l} \times \sum_{j=1}^{n} Y_{ij\gamma\beta}^{2}}{m\,\ln}\right)^{2}}.$$

The value of measured daily schedule load can be given as a matrix-current or power, where:

i : rows (at half-hour intervals $i = 48$ per day, at hourly intervals $i = 24$).

j : columns of the matrix (at 5-minute increments of the load curve in a half-hour interval $j = 7$, $j = 13$ in an hour interval).

Investigation of the electrical load of public buildings has its own peculiarity. Average maximum values of different groups of public buildings cannot be subjected to similar statistical processing as the load of residential buildings, so there are large variations in the average value of apartments load. Because absolute values of electrical loads for different by construction and character of public building have a very large variation of average values and do not provide sufficient accuracy of calculations. Therefore, it is necessary to conduct processing of public buildings statistical data for homogeneous consumers.

4 The Algorithm of a New Method Implementation

The algorithm of the new method is as follows [4]:

- The experimental investigations of urban power consumers and electrical loads conduct.
- Using the mathematical apparatus of the theory of probability and mathematical statistics are being built characteristic (model) daily schedules of electrical loads for all types of urban power consumers and the electrical load.
- On the basis of these daily schedule characteristic, using Lagrange, Newton, Stirling, Bessel, Chebyshev and others methods are determined analytical functions (equations, formulas) of curves (daily schedules) of mode changes each consumer.
- Next, by functions (equations, formulas), which are describes curve of the daily schedule load, analytically are defined of maximum coefficients Km.

In general, the equation describing the daily schedule of electric load has the form:

$$S\left(x\right) = U_{i=1,\cdots,24} S_i\left(x\right),$$

or

$$\begin{aligned}
S_3\left(x\right) &= \left(x_i^3 - 3x_i^2 x + 3x_i x^2 - x^3\right) + \left(x^3 - 3x^2 x_{i-1} + 3x x_{i-1}^2 - x_{i-1}^3\right) + \frac{f_{i-1}}{h_i} x_i \\
&\quad - \frac{M_{i-1}\cdot h_i^2}{6h_i} x_i - \frac{f_{i-1}}{h_i} x + \frac{M_{i-1}\cdot h_i^2}{6h_i} x + \frac{f_i}{h_i} x - \frac{f_i}{h_i} x_{i-1} - \frac{M_i \cdot h_i^2}{6h_i} x + \frac{M_i \cdot h_i^2}{6h_i} x_{i-1} \\
&= (M_i - M_{i-1}) \cdot \frac{x^3}{6h_i} + \frac{3}{6h_i} \cdot (M_i x_i - M_{i-1} x_{i-1}) x^2 + (M_i \cdot {x^2}_{i-1} - M_{i-1} \cdot {x^2}_i - 2f_{i-1}, \\
&\quad + M_{i-1} \cdot {h^2}_i - M_i \cdot {h^2}_i + 2f_i) x + (M_{i-1} \cdot {x^3}_i - M_i \cdot {x^3}_{i-1} + 6f_{i-1} \cdot x_i - 6f_i \cdot x_{i-1} \\
&\quad - M_{i-1} \cdot h_i^2 \cdot x_i + M_i \cdot h_i^2 \cdot x_{i-1}) \frac{1}{6h_i} \\
&= a_{i0} \cdot x^3 + a_{i1} \cdot x^2 + a_{i2} \cdot x + a_{i3},
\end{aligned}$$

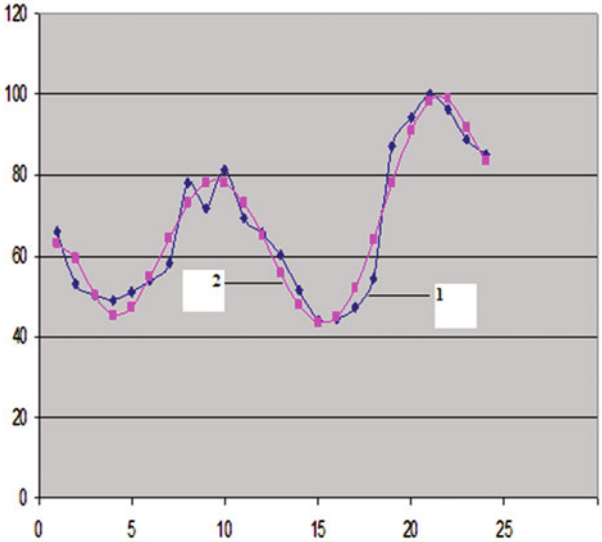


Fig. 3. The characteristic (typical) daily schedule of electric load of the shop manufactured goods: 1- constructed from experimental data, 2- built by Lagrange polynomial

where

$$
\begin{aligned}
a_0 &= \frac{M_i - M_{i-1}}{6h_i}, \\
a_1 &= \frac{1}{2h_i}(M_{i-1}x_i - M_i \cdot x_{i-1}), \\
a_2 &= \frac{1}{2h_i}(M_i \cdot {x^2}_{i-1} - M_{i-1} \cdot {x^2}_i + M_{i-1} \cdot {h^2}_i - M_i \cdot {h^2}_i - 2f_{i-1} + 2f_i), \\
a_3 &= \frac{1}{6h_i}(M_{i-1} \cdot {x^3}_i - M_i \cdot {x^3}_{i-1} + 6f_{i-1} \cdot x_i - 6f_i \cdot x_{i-1} \\
&\quad + M_i \cdot h_i^2 \cdot x_{i-1} - M_{i-1} \cdot h_i^2 \cdot x_i).
\end{aligned}
$$

On the base of this algorithm, substituting the values M_i, f_i, h_i, x_i from the table, we obtain a system from 24 equations in 3rd order.

As a result, for the curve of the daily schedule of electric load we obtain a system of equations, which has the following general form:

$$S_3\,(x) = a_{i0}x^3 + a_{i1}x^2 + a_{i2}x + a_{i3},$$

where a_{i0}, a_{i1}, a_{i2}, a_{i3}: coefficients of each equation.

Substituting the data from the daily schedule in these equations, we can obtain the curve segments of the daily schedule, and combining all of the 24 segments, as a result, get the curve of daily schedule, that is almost completely coinciding with the characteristic of daily schedule of consumer loads.

For each electrical consumer or appliances construct its equation of the curve and own formula for the maximum electrical load. As a result, the calculation is obtained precise and more authentic and closer to the actual load, taking into account the peculiarities of load conditions change, the time factor for each appliance and consumer included in the city's power system.

Example 1. The characteristic (typical) daily schedule of electric load of shop manufactured goods, built using the Lagrange polynomial for $x_0 = 1$, $x_1 = 10$, $x_2 = 16$, $x_3 = 22$; $y_0 = 10$, $y_1 = 54$, $98 = v_2$, $v_3 = 10$ is as follows (Fig. 3) [4]:

$$P\left(t\right) = 21.95057 - 15.91852t + 2.72962t^2 - 0.26084t^3.$$

5 Conclusions

The distributed generators of electricity integrated into Unified Energy System (UES), represent difficult system in which communication between observable parameters and a system condition has difficult and ambiguous character. Therefore, the approach to management in the automatic control systems, based on management on output parameters, is inadequate as a matter of fact. In this case formation of operating influence should be carried out on the basis of data about a current condition of object of management.

The daily schedules of electric load allow properly assess the mode of the work of electrical equipment and the enterprise as a whole, identify the bottlenecks and reserves, to set the optimal mode of the work.

This paper develops an improved method for determining the estimated electrical loads in residential and public buildings with using characteristic of daily load curves of the same type of consumers, constructed from experimental measurements of the characteristics of the typical power consumption of residential and public buildings consumers.

The use of the developed methods, models and algorithms allow more reliably determine the electrical loads in modern residential and public buildings taking into account peculiarities the load condition change each electrical consumer and load, included in the city's power supply system.

References

1. Arifjanov AS, Ayupov AS (2015) Forecasting and rationalization of energy consumption for megacities taking into account the factors caused by scientific and technical progress, with application of informational - analytical technologies. In: A report on the global forum for energy, environment and commercial civilization (GFEECC)
2. Arifjanov AS, Zakhidov RA (2017) An approach to the creation of the adaptive control system for integration of nonsteady power sources into a common electric power grid. In: Proceedings of the tenth international conference on management science and engineering management. Springer, pp 563–574

3. Arifjanov AS, Zakhidov RA, Saidkhodjaev AG (2014) Determination of electrical loads in modern residential and public buildings. In: First all-Russian scientific-practical conference “Energy efficiency and energy saving: theory and practice”
4. Arifjanov AS, Yusupov ZE, Saidkhodjaev AG (2016) The methods, models and algorithms of electrical loads estimation for different type’s electricity consumers in the residential and public buildings. In: Proceedings of the 3rd international symposium on environment and morality, SEM 2016
5. Saidkhodjaev AG, Teshabayev BM (2010) A method for determining the maximum load of electric network and the maximum electrical load of individual consumers
6. Volobrinsky SD, Kayal GM et al (1971) The electrical load of the industrial enterprises. Leningrad branch
7. Zakhidov RA (2015) Development of renewable energy and implementation of programs of energy saving in Uzbekistan as a basis of ensuring power safety. In: A report on the global forum for energy, environment and commercial civilization (GFEECC)

Haze-Related Air Pollution and Impact on the Stock Returns of the Listed Steel Companies in China

Kai Liu[1], Ying Li[1](✉), and John Thomas Delaney[2]

[1] Business School, Sichuan University,
Chengdu 610064, People's Republic of China
liyinggs@scu.edu.cn

[2] Kogod School of Business, American University,
4400 Massachusetts Ave, Washington, D.C. 20016-8044, USA

Abstract. The purpose of this paper is to study the effects of air pollution, especially the haze, on the stock returns of steel mills. This study collects data from air quality index, variables represent characteristic of the eleven steel mills in China and the stock returns ratio from the stock market etc. SPSS19.0 is utilized to conduct a descriptive analysis of the correlation between the principal air pollution(including $PM_{2.5}$, PM_{10}), the variables represent charactering(monetary funds, net assets, liabilities, operating margin, financial leverage, and total asset growth) of the elven steel mills in China and the stock returns ratio from the stock market etc. Hence, through the research on the air pollution index, analysis of the listed company's earnings and stock price, and by using the linear regression analysis method, research shows that the serious air pollution have a negative impact on the profitability of iron and steel enterprises through the emotion and expectations of investors. It is imperative for people to tackle air pollution urgently.

Keywords: Haze related air pollution · Stock return rate · Listed steel companies

1 Introduction

With the rapid economic development and fast expansion of productive enterprises, China is increasingly facing the large-scale severe air pollution. Air pollution with $PM_{2.5}$ and PM_{10} as the main pollutants troubles the people's life, threatening their health. According to a report released by the WHO in May, 2016, seven of the world's 10 cities with the most severe air pollution are in China (China Environmental Report 2016). $PM_{2.5}$ has an essential impact on air quality and visibility, and the haze may last longer based on the different geographical and meteorological conditions, impairing the people' health considerably. The report further shows that less than 1% of China's 500 largest cities

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_100

meet the air quality standards recommended by the WHO (China Environment Report 2016).

The severe air pollution is caused principally by the rapid growth of productive enterprises which use coal as the main raw material and highlight high energy consumption. To maximize the profits, these steel companies have ignored environmental protection, and even sacrificed the environment for a long time, causing serious environmental problems, threatening future health of the Chinese people and sustainable development of the ecosystem. As a resource-intensive industry, the iron and steel industry, using coal as the main source of energy, explores a variety of powdery and massive ferrous metals and non-metallic minerals through large-scale production and complex processes. The emissions of the iron and steel industry can be divided into three categories: The first is the exhaust gas caused in the production process, such as the smoke, sulfur dioxide and other harmful gases generated through the sintering, smelting, and steel rolling process. The second type is the dust and sulfur dioxide produced by burning fuels, for instance, coal and coal gas, in the furnace. The third category is the dust generated during the processes of transporting, loading and discharging and processing of raw materials and fuels.

The smog not only severely impairs people's health but limits their outdoor activities as well. It affects individual mental activities and emotional states, resulting in more negative emotions. This paper seeks to study whether air pollution, especially the smog, affects the steel plants' stock returns. Through examining the air pollution index, the paper analyzes the earnings and stock prices of listed steel companies, and uses the linear regression analysis to investigate whether the air pollution has relevance to the stock returns of these enterprises. The structure of this paper is as follows: Introduction comes the first, which is followed by a literature review–the second section. The third part is the research model, data and the method. Then, the paper presents the empirical research results and the corresponding analyses and discussions, which constitute the fourth section. Last, it draws conclusions and provides suggestions.

2 Literature Review

Studies have shown that air pollution directly or indirectly affects people's psychological conditions and emotions. Lepori [8], using different trading techniques in Italy and samples from major international stock exchanges, demonstrated the relationship between air pollution, people's emotions, as well as stock returns. Evans [5] find that exposure to polluted air increases the level of depression, anxiety, helplessness, and anger. Emotion affects people's assessment of both future prospects and risks [13]. Some studies reveal that air pollution is positively correlated with people's negative emotions. Levy [9] examined the relationship between air pollution and stock returns. Using air quality indices and stock return data from four US stock exchanges, they argue that air pollution is negatively correlated with stock returns. Mehra [11] observed, in their study, that small emotional volatility has significant impacts on the fluctuation of capital prices. Lucy [10] held that emotions, which are not caused by the investors' future decisions, may still affect the decision makers.

To address the deteriorating air quality and ease the increasing public concern, relevant government departments may reinforce stringent emission standards and formulate strict environmental policies to control the high pollution and high emissions behavior. The regulatory policies strengthening environmental protection will affect the stock returns of iron and steel companies with the such characteristics as high input, high pollution, and high energy consumption [12]. Taking China's steel mills as examples, considering the serious environmental protection attitudes and stringent environmental standards, steel mills will face greater pressure from environmental protection and environmental costs, and the stock returns of the listed steel companies will be affected accordingly. Meanwhile, air pollution may easily trigger industrial policy adjustments, which, in turn, lead to the re-allocation of resources. This may impact the listed companies' operating performances and growth prospects, resulting in stock price fluctuation [4]. In addition, other government policies, for instance, security regulations, energy prices, credit rationing, and international cooperation, which are closely related to air quality, also have a significant impact on the stock prices of the listed companies [6].

It is worth noting that air quality is regional and it largely affects the sentiment of local investors, especially individual investors. Likewise, air pollution may impact the stock market through influencing the stock traders' emotion. Currently, the two major stock exchanges and three futures exchanges in China adopt the command-driven system rather than the quotation-driven system, and all trading quotations should be input to the matching system by the investors with the help of the agent brokers (exchange members). As for the command trading system operators, when their emotions are affected by local air pollution, their rational judgment and selection ability may be reduced, resulting in irrational trading behavior and causing stock price fluctuation. In the information age, individual investors are still the largest fund providers and the main traders in the current stock market [3], among whom there is a "herd effect" [1,2,7]. In this context, when outside investors make trade decisions (of the listed companies) based on the air quality of the major cities (e.g. Beijing, Shanghai), the "herd effect" may further amplify the impact of air quality on the stock market of the steel industry.

3 Model and Assumption

3.1 Data

This paper studies 11 Chinese steel plants, which are Shanghai Baosteel (600019), Shougang Group (000959), Ansteel Group Corporation (000898), Wuhan Iron and Steel (Group) Company (600005), Jiangsu Shagang Group Co., Ltd. (0020), Taiyuan Iron & Steel (Group) Co., Ltd. (000825), Jinan Iron and

Steel Co., Ltd. (600022), Panzhihua Iron and Steel (Group) Co. (000629), Hesteel Group Tangsteel Company (000709), Benxi Iron and Steel (000761), and Nanjing Iron and Steel Group Co., Ltd. (600282). The paper gathers the stock returns and feature variables (monetary funds, net assets, liabilities, operating margin, financial leverage, and total asset growth) from the CSMAR database. Moreover, some data of financial leverage and total asset growth are collected from the WIND database and China Securities Depository and Clearing Corporation Limited (CSDC).

This paper uses six variables (monetary funds, net assets, liabilities, operating margin, financial leverage, and total asset growth), which are borrowed from Berger and Tuttle, to examine the stock information of the steel companies from different dimensions. The annual financial statements are released with a time lag. That is, the announcements are usually released in March of the next year. In order to ensure the relative effectiveness of the financial information, the paper uses the financial information (July, t, $\sim$ June, $t+1$) as the stock transaction data of the $t-1$ year.

Table 1. The definitions of explained variables

Explained variables	Comments
Monetary funds	Indicates the size of the company
Net profit assets	Indicates the company's ability to develop
Debt ratio	Indicates the company's ability to repay the debt
Operating margin	Indicates the company's growth characteristics
Financial leverage	Indicates the company's profitability
Total assets growth rate	To determine the stability of the stock return rate

The urban air quality data used in this study are gathered at the sites of the steel plants. Partly of the data are from the $PM_{2.5}$ hourly data monitored by the five US embassies and Consulates in China. The 2013–2015 air quality index data (such as SO_2, NO_2, CO, $PM_{2.5}$, PM_{10}, and AQI) are obtained from the Ministry of Environmental Protection of the People's Republic of China.

3.2 The Model

In order to investigate the impact of air pollution on the earnings of the steel mills, the paper uses an ordinary least squares (OLS) regression model, which is as follows:

$$\text{For Returns} = \alpha^{X_1} + \beta^{X_2} + e.$$

For Returns represents the dependent variable and refers to the monthly rate of return of the steel plant.

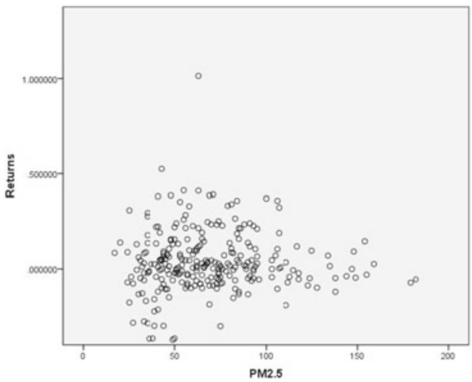


Fig. 1. A scatter plot of returns and $PM_{2.5}$

In the model, α denotes a constant term; e is the random error term; β signifies the weight associated with the explanatory variable; X_1 is $PM_{2.5}$ and X_2 represents PM_{10}.

The purpose of this paper is to explain whether and how air quality affects the profitability of the iron and steel plants. The explanatory variables involved are mainly the air quality indicators. The explained variables include data such as monetary funds, net profit assets, debt ratio, operating margin, financial leverage and total assets growth rate (Table 2).

The linear model: $Y = -0.179 \times X_1 + 0.233 \times X_2 + e$, where Y represents the returns, X_1 is $PM_{2.5}$, X_2 denotes PM_{10}, and e signifies the error.

In Table 3, the correlation coefficient R^2 is 0.015, which is very small and close to 0. Therefore, the returns are not significantly correlated to $PM_{2.5}$ and PM_{10}, and the e of the model is larger.

According to the analysis of the above theoretical background, in this paper, the assumption is as follows serious air pollution have a negative impact on the profitability of iron and steel enterprises.

Table 2. Model summary

Model	R	R Square	Adjusted R Square	Std. Error of the estimate	Change statistics				
					R Square change	Change F	df1	df2	Sig. F change
1	0.122a	0.015	0.007	0.160499	0.015	1.972	2	260	0.141

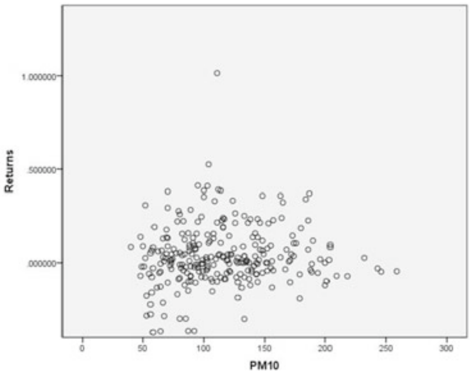


Fig. 2. a scatter plot of returns and PM_{10}

Table 3. Coefficient analysis

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.	Correlations			Collinearity statistics	
		B	Std. error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	0.003	0.028		0.123	0.903					
	$PM_{2.5}$	−0.001	0.001	−0.179	−1.505	0.133	0.02	−0.093	−0.093	0.267	3.749
	PM_{10}	0.001	0	0.233	1.958	0.051	0.08	0.121	0.121	0.267	3.749

4 Empirical Results and Analysis

4.1 Descriptive Analysis

General descriptive statistical analysis & descriptive statistical analysis by company names (Table 4).

Table 4. The overall descriptive statistics are summarized

	N	Min	Max	Mean	Std. deviation	Variance	Skewness	Std. error of skewness	Kurtosis	Std. Error of kurtosis
Returns	263	−0.3729	1.0142	0.0392	0.1611	0.026	1.006	0.15	5.056	0.299
$PM_{2.5}$	275	17	182	67.95	30.103	906.212	1.072	0.147	1.311	0.293
PM_{10}	275	40	259	113.84	42.381	1796.152	0.729	0.147	0.241	0.293
Valid N (list wise)	263									

The population standard deviation of net profit is very large, indicating significant fluctuations of the net profit. The variance of $PM_{2.5}$ and PM_{10} is relatively small, but also they are over 30, reflecting obvious fluctuations of the air quality. The skewness of $PM_{2.5}$ and PM_{10} is greater than 0 and has a maximum, which should be removed, suggesting the right deviation. The skewness of the net profit is less than 0, showing left deviation. The skewness has a minimum, which should be removed (Table 5).

Table 5. The conclusion and analysis of the overall descriptive statistics

No		N	Min	Max	Mean	Std. deviation	Variance	Skewness	Std. Error of skewness	Kurtosis	Std. Error of kurtosis
1	Returns	25	−0.2231	0.3908	0.0193	0.1249	0.016	0.958	0.464	2.783	0.902
	$PM_{2.5}$	25	33	124	55.56	19.812	392.507	1.865	0.464	4.881	0.902
	PM_{10}	25	49	‘	75.64	22.381	500.907	1.143	0.464	1.176	0.902
	Valid N	25									
	(listwise)	25									
2	Returns	21	−0.3728	0.2349	0.0420	0.1466	0.022	−0.832	0.501	1.706	0.972
	$PM_{2.5}$	25	44	154	80.84	28.868	833.39	1.19	0.464	1.043	0.902
	PM_{10}	25	58	173	108.28	30.893	954.377	0.275	0.464	−0.774	0.902
	Valid N	21									
	(listwise)	21									
3	Returns	25	−0.2137	0.4106	0.0230	0.1227	0.015	1.075	0.464	3.175	0.902
	$PM_{2.5}$	25	40	144	64.64	24.605	605.407	1.609	0.464	3.206	0.902
	PM_{10}	25	66	188	109.72	29.389	863.71	0.736	0.464	0.576	0.902
	Valid N	25									
	(listwise)	25									
4	Returns	25	−0.2859	0.3299	0.0257	0.1435	0.021	0.323	0.464	0.131	0.902
	$PM_{2.5}$	25	35	182	78.44	38.448	1478.257	1.523	0.464	2.355	0.902
	PM_{10}	25	67	218	113.08	37.627	1415.827	1.214	0.464	1.502	0.902
	Valid N	25									
	(listwise)	25									
5	Returns	19	−0.1911	1.0142	0.1490	0.2862	0.082	1.622	0.524	3.538	1.014
	$PM_{2.5}$	25	33	111	63.12	22.769	518.443	0.728	0.464	−0.191	0.902
	PM_{10}	25	66	179	113.04	27.902	778.54	0.658	0.464	0.022	0.902
	Valid N	19									
	(listwise)	19									
6	Returns	24	−0.3657	0.3861	0.0370	0.1712	0.029	−0.192	0.472	1.102	0.918
	$PM_{2.5}$	25	38	105	64.72	21.084	444.543	0.562	0.464	−1.091	0.902
	PM_{10}	25	82	176	122.12	27.039	731.11	0.123	0.464	−0.863	0.902
	Valid N	24									
	(listwise)	24									
7	Returns	25	−0.3012	0.3698	0.0337	0.1503	0.023	0.447	0.464	0.851	0.902
	$PM_{2.5}$	25	57	159	91.68	26.009	676.477	1.326	0.464	1.194	0.902
	PM_{10}	25	98	259	173.4	38.249	1463	0.414	0.464	0.157	0.902
	Valid N	25									
	(listwise)	25									
8	Returns	25	−0.2836	0.3081	0.0278	0.1369	0.019	0.121	0.464	0.439	0.902
	$PM_{2.5}$	25	17	73	35.72	13.296	176.793	1.233	0.464	1.615	0.902
	PM_{10}	25	40	144	74.56	26.013	676.673	1.182	0.464	1.261	0.902
	Valid N	25									
	(listwise)	25									
9	Returns	24	−0.3657	0.3861	0.0370	0.1712	0.029	−0.192	0.472	1.102	0.918
	$PM_{2.5}$	25	39	147	92.04	29.464	868.123	0.211	0.464	−0.82	0.902
	PM_{10}	25	80	204	153.68	37.201	1383.893	−0.393	0.464	−0.962	0.902
	Valid N	24									
	(listwise)	24									
10	Returns	25	−0.2762	0.3809	0.0324	0.1365	0.019	0.357	0.464	1.303	0.902
	$PM_{2.5}$	25	26	111	53.32	22.527	507.477	0.791	0.464	−0.108	0.902
	PM_{10}	25	55	146	94.08	27.412	751.41	0.326	0.464	−0.935	0.902
	Valid N	25									
	(listwise)	25									
11	Returns	25	−0.3672	0.4124	0.0312	0.1607	0.026	0.067	0.464	0.942	0.902
	$PM_{2.5}$	25	30	155	67.32	29.22	853.81	1.387	0.464	2.432	0.902
	PM_{10}	25	59	243	114.6	44.507	1980.917	1.188	0.464	1.735	0.902

4.2 Descriptive Statistical Analysis by Companies

In terms of companies, the average net income of Panzhihua Iron and Steel (Group) Co. is negative, suffering loss, while the average $PM_{2.5}$ and PM_{10} of this enterprise are 35.72 and 74.56, respectively, considerably lower than the overall $PM_{2.5}$ and PM_{10} average, namely 67.95 and 113.84. It shows that the air quality of Panzhihua Iron and Steel (Group) Co. is better than the overall situation. The Shanghai Baosteel Group has the highest net income, the average $PM_{2.5}$ and PM_{10} of which are 55.56 and 75.64, lower than the overall $PM_{2.5}$ and PM_{10} average. As for other data, the net income fluctuates substantially, and the $PM_{2.5}$ and PM_{10} are close to the overall mean. Therefore, the relationship between $PM_{2.5}$ and PM_{10} and steel plants' net income cannot be analyzed directly (Table 6).

Table 6. Descriptive statistical test: Chi-Square tests

Test of $PM_{2.5}$	Value	df	Asymp. sig.(2-sided)
Pearson Chi-Square	23098.601[a]	23100	0.501
Likelihood ratio	2243.815	23100	1
Linear-by-linear association	0.109	1	0.741
N of valid cases	263		
Test of PM_{10}	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	30356.775[a]	30492	0.707
Likelihood ratio	2395.163	30492	1
Linear-by-linear association	1.666	1	0.197
N of valid cases	263		

[a] 30856 cells (100.0%) have expected count less than 5. The minimum expected count is .00.
[b] 23432 cells (100.0%) have expected count less than 5. The minimum expected count is .00.

The Table 1 is a chi-square test of $PM_{2.5}$, while the second is a chi-square test of PM_{10}. It can be seen from the tables that $sig < 0.05$, meaning that there are significant differences between $PM_{2.5}$, PM_{10} and the net income (Table 7).

The Table 1 is about $PM_{2.5}$ and the second is about PM_{10}. The values of the symmetric measures are both around 0.99, which is much larger than 0.1, indicating correlation between $PM_{2.5}$, PM_{10}, and the net income.

4.3 Correlation Analysis

As shown in the PM_{10} and net profit correlation Table 8, for most areas, the absolute value of the Pearson correlation coefficient between the net profit and PM_{10} is relatively low, suggesting an insignificant impact of PM_{10} on the net profit. Only a few areas are excepted: For instance, Jiangsu shows strong correlation.

Table 7. Symmetric measures

Test of $PM_{2.5}$		Value	Asymp. Std. error[a]	Approx. T[b]	Approx. sig
Nominal by nominal	Phi	9.372			0.501
	Cramer's V	0.937			0.501
	Contingency coefficient	0.994			0.501
Ordinal by ordinal	Kendall's tau-b	0.022	0.042	0.534	0.593
	Kendall's tau-c	0.022	0.042	0.534	0.593
	Gamma	0.022	0.042	0.534	0.593
	Spearman correlation	0.036	0.062	0.575	.566[c]
Interval by interval	Pearson's R	0.02	0.051	0.33	.742[c]
N of valid cases		263			
Test of PM_{10}		Value	Asymp. Std. error[a]	Approx. T[b]	Approx. sig
Nominal by nominal	Phi	10.744			0.707
	Cramer's V	0.935			0.707
	Contingency coefficient	0.996			0.707
Ordinal by ordinal	Kendall's tau-b	0.055	0.041	1.328	0.184
	Kendall's tau-c	0.055	0.041	1.328	0.184
	Gamma	0.055	0.041	1.328	0.184
	Spearman correlation	0.083	0.062	1.341	.181[c]
Interval by interval	Pearson's R	0.08	0.053	1.292	.197[c]
N of valid cases		263			

[a]Not assuming the null hypothesis.
[b]Using the asymptotic standard error assuming the null hypothesis.
[c]Based on normal approximation.

This correlation varies from region to region. For example, Shanghai and Jiangsu show a positive correlation. That is, the net profit increases with the rise of PM_{10}. In contrast, Nanjing shows a strong negative correlation. In other words, the net profit reduces with an increase in PM. Overall, the majority of the cities show a negative correlation. That is, when the PM_{10} increases, the net profit decreases.

According to the correlation analysis of PM_{10} and the total assets growth, different from the net growth rate, most regions show strong positive correlations. That is, the total asset ratio will rise with the increase in PM_{10}. For example, Jinan, Panzhihua, Tangshan, and Benxi show a strong positive correlation. Jiangsu, however, shows a strong negative correlation.

The correlation coefficient between $PM_{2.5}$ and the return is less than 0.2, reflecting a weak correlation. Although some correlation coefficient of PM_{10} is over 0.2, but they are less than 0.5, indicating its weak correlation with the return. $PM_{2.5}$ and PM_{10}, hence, impact the return of the steel companies.

Table 8. Correlations

		Returns	$PM_{2.5}$	PM_{10}
Pearson correlation		1	0.02	0.08
	$PM_{2.5}$	0.02	1	0.856
	PM_{10}	0.08	0.856	1
Sig. (1-tailed)	Returns	.	0.371	0.099
	$PM_{2.5}$	0.371	.	0
	PM_{10}	0.099	0	.
N	Returns	263	263	263
	$PM_{2.5}$	263	263	263
	PM_{10}	263	263	263
		Returns	$PM_{2.5}$	PM_{10}
Returns	Pearson correlation	1	0.02	0.08
	Sig. (2-tailed)		0.742	0.197
	N	263	263	263
$PM_{2.5}$	Pearson correlation	0.02	1	.857[a]
	Sig. (2-tailed)	0.742		0
	N	263	275	275
PM_{10}	Pearson correlation	0.08	.857[a]	1
	Sig. (2-tailed)	0.197	0	
	N	263	275	275

[a] Correlation is significant at the 0.01 level (2-tailed).

5 Conclusion and Suggestion

This paper, based on a theoretical analysis, argues that air quality may affect stock market participants through emotions, policies, and expectations, ultimately impacting the steel companies' rate of return.

The results of this study show that severe air pollution has a negative effect on the profitability of the steel companies. The air pollution is not only an environmental problem, but also an economic and financial issue. Therefore, it is imperative for people to tackle air pollution urgently. The steel companies should increase the public transparency of information, enhance the development of corporate social responsibility activities, improve the corporate image. Air pollution event occurs, investors should keep rational judgment and choice ability, make relevant rational trading behavior. Relevant government departments should reinforce stringent emission standards and formulate strict environmental policies to control the high pollution and high emissions behavior to address the deteriorating air quality and ease the increasing public concern.

The research enriches the literature regarding the influence of investor sentiment on the capital market, and provides the empirical estimation of the impact of air pollution on the Chinese capital market. It has certain reference value for the relevant government departments in terms of policy evaluation and decision-making.

Acknowledgements. This research was supported by System Science and Enterprise Development Research Center Project (Grant No. Xq16C12), Sichuan University Central University Research Fund project (Grant No. skqy201653), Soft Science Project of Sichuan Science and Technology Department (No. 2017ZR0033) and Sichuan Social Sciences Research 13th Five-Year annual planning project, (NO. SC16A006).

References

1. Banerjee AV (1992) A simple model of herd behavior. Q J Econ 107(3):797–817
2. Cao M, Wei J (2005) Stock market returns: a note on temperature anomaly. J Bank Finance 29(6):1559–1573
3. Chen YA, Tan S (2005) Estimation of the steady inflation rate of economic growth. Econ Res J 4:002
4. Esposito P, Patriarca F et al (2013) Economic convergence with divergence in environmental quality? Desertification risk and the economic structure of a mediterranean country (1960–2010). Mpra Paper 102(4):715–721
5. Evans GW, Jacobs SV et al (1987) The interaction of stressful life events and chronic strains on community mental health. Am J Community Psychology 15(1):23–34
6. Fong WM, Toh B (2014) Investor sentiment and the max effect. J Bank Finance 46(3):190–201
7. Kamstra MJ, Kramer LA, Levi MD (2003) Winter blues: a sad stock market cycle. Am Econ Rev 93(1):324–343
8. Lepori GM (2015) Air pollution and stock returns: Evidence from a natural experiment. J Empir Finance 35:25–42
9. Levy T, Yagil J (2011) Air pollution and stock returns in the US. J Econ Psychology 32(3):374–383
10. Lucey BM, Dowling M (2005) The role of feelings in investor decision-making. J EconSurv 19(2):211–237
11. Mehra R, Sah R (2002) Mood fluctuations, projection bias, and volatility of equity prices. J Econ Dynamics Control 26(5):869–887
12. Oberndorfer U (2006) Environmentally oriented energy policy and stock returns: an empirical analysis. ZEW-Centre for European Economic Research Discussion (06–079)
13. Slovic P, Finucane ML et al (2007) The affect heuristic. Eur J Oper Res 177(3):1333–1352

Low Carbon-Oriented Coupling Mechanism and Coordination Model for Energy Industry Investment and Financing of China

Yanfei Deng[1] and Lei Xu[2(✉)] and Ning Hou[3]

[1] School of Management, Southwest University for Nationalities, Chengdu 610064, People's Republic of China
[2] School of Economics, Xihua University, Chengdu 610039, People's Republic of China
leihsu@163.com
[3] Faculty of Engineering, Imperial College London, London, UK

Abstract. Based on the perspective of low carbon and greenhouse gas emission reduction, this paper utilizes the synergy theory to analyze the coupling and coordination characteristics of China's energy investment and financing, and argues that energy investment and energy financing constitute a composite system and find that it has a stress effect on the climate environment, and the climate environment has a constraint effect on energy investment and financing. According to the analysis of the coupling relationship between the two, the energy industry investment and financing coupling coordination model and evaluation index system are established. By revealing the development law and evolution trend between the two, we will actively promote the coordinated and coordinated development of economic growth, energy supply and climate and environment, and provide new research ideas and analysis methods.

Keywords: Renewable energy · Investment and financing · Coupling mechanism · Coordination model

1 Introduction

In recent years, affected by the combined effects of resource endowment constraints and economic growth, China's energy production and energy consumption is dominated by coal, greenhouse gas emissions are rising rapidly has become the first emissions of global carbon dioxide and sulfur dioxide [5]. In recent years, China's atmospheric pollutant emissions continue to increase, the composite air pollution has become increasingly prominent, sulfur dioxide, nitrogen oxides and volatile organic compounds led to the second pollution was intensified, which include Fine particulate matter, ozone, acid rain and so on. The formation cause of air pollution is complex energy production and processing conversion is one of the important source. So it is imperative to realize energy saving and emission

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_101

reduction and low-carbon development in the energy industry. Energy industry investment and financing is to regulate and control the energy industry to improve the level of its clean production, “the total valve”, both the coupling coordinated development will help to accelerate the existing technology and equipment to upgrade and clean production projects and technology operation, to fundamentally reduce the energy consumption, reduce industrial emissions, to promote the coordinated development of energy, economy and environment, the purpose of climate. Investment in energy industry and energy industry financing interaction coupling relationship as follows: Energy industry investment through the production of investment and pollution control investment in two aspects of stimulating financing development, energy industry financing through the support of the above two aspects to promote investment expansion. When the energy industry investment expansion too fast and beyond the ecological environment carrying capacity limit, the government based on the reality of environmental degradation will significantly enhance the policy constraints of environmental protection and greenhouse gas emission reduction. Under the comprehensive control of industrial policy, credit policy and environmental protection policy, the energy industry financing will continue to shrink and restrict the energy industry investment. Instead, consistent with the national industrial policy, climate and environment protection policy of renewable energy projects in the production, energy production source pollution control and cleaner production projects will receive financing inclination, thereby stimulating investment in areas related to energy industry development. When the pollution control of energy production of excessive growth of investment, the economic principle of optimal use of funds is not consistent, the development of energy industry investment will be blocked due to the lack of funds, climate and ecological environment protection also lost the support of the construction. With the rapid growth of economy, the increasingly intensified climate environment has become an important bottleneck in the development of energy economy. This paper is based on the macro policy background of greenhouse gas emission reduction from the theory of synergetic and coupled coordination evaluation method the energy industry investment and energy industry financing can be considered as two subsystems, in the analysis of the two coupling based on the coordination of the relationship, the establishment of energy industry investment and financing coupled coordination model and the order parameter evaluation index system. By revealing the law of development and evolution trend of the two, to provide a new research ideas and analysis methods for the effective treatment of economic growth energy supply and climate environment deterioration of the prominent contradictions and actively promote the coordinated development of coupling.

Coupling coordination theory is mainly used to study economic management in technology innovation and technology management, industry cluster and industry chain and ecological environment and regional economic problems [1,3,6,7]. American scholar Weick [8] first used coupling theory to study economic and social problems, he explained the relationship between the school members contact each other but keep each other independent through loose coupling theory. Norgaad [4] put forward the theory of coordinated development, think through the feedback loop and can achieve the coordination and common development between social system and ecological system. Guneralp [2] constructed the important infrastructure and the national security, the ecological environment, the economical coupling effect model and so on. The domestic scholars based on the measure method of coupling coordination, researches on the problems of ecological environment and economic development coordination. Binding energy industry investment and financing characteristics, we believe that they may exist coupling interaction and coordinated operation mechanism and the coupling coordination theory and method shall also apply in the energy industry investment financing in the field of research on the coordination mechanism. In the plane of the sustainable development, if the interaction of investment in the energy industry and the energy industry financing has a higher level of coordination of the coupling of, there will be helping to improve the efficiency of the compound system, and ultimately the economy energy low carbonization and intensive development.

2 Key Problem Statement

Coupling as a physical concept, the extent to which two (or more than two) systems or forms of motion are affected by each other. For example, between two connected to a spring pendulum, they shake each other ups and downs, mutual influence, this interaction is called coupling pendulum. Similarly, you can put two sub investments in energy industry and energy industry financing system through their respective coupling elements produce interaction and mutual influence of the phenomenon is defined for the coupling of the investment in the energy industry and the energy industry financing, both of which constitute the energy industry investment and financing system. Energy industry investment subsystem as the guide and engine of energy economic development, determines the quality, scale and benefit of the energy industry financing subsystem, and at the same time, it is influenced by the movement of the order parameter of the financing of the energy industry. The mutual coupling situation of the investment and financing of the energy industry is decided by the investment and financing gap, which is equal to the time accumulation of investment minus the financing. Any investment factors that affect the investment and financing gap will contribute to the change of the financing. Under the background of greenhouse gas emissions reduction, investment in the energy industry in addition to the category of traditional energy industry investment should also cover for prevention and control of pollution source, newly started construction projects "three simultaneous" project, renewable energy development, carbon capture and sequestration (CCS) and integrated gasification combined cycle technology (IGCC) and other new production projects and processing technology of the investment. The productive capacity of these investments will contribute to a significant reduction in industrial emissions and greenhouse gases, and improve the quality of regional climate and environment. Information related to investment will continue to feedback to departments of environment and trade market, through policy intervention and price signals to attract financing support, and achieving the energy industry investment and financing coupling coordinated development goals.

In the energy industry investment subsystem, the long-term implementation of the energy industry investment government approval, corporate finance, bank loans investment and financing model, the lack of flexible, effective and clean investment and financing mechanism. This type of investment is the subject of the behavior of state-owned energy enterprises, national investment of this kind of enterprises lack of supervision method and risk control means, lack of supervision and evaluation mechanism of enterprise investment decision, enterprise project approval lack of benefit and risk evaluation mechanism. Especially in the oil and gas, electric power and other industries, to a large extent there are administrative intervention, unclear definition of the functions of the government, macro-control offside, resulting in energy companies are often blind investment behavior. In the energy industry financing subsystem, in recent years, the energy industry financing channels tend to be diversified, but still in the enterprise internal self financing and bank credit as the main. Commercial banks in the energy companies monopoly profits and national climate and environmental protection policies to make a difficult choice between. Leading to the ups and downs of the size of the finance. As shown in Fig. 1, the energy industry investment and financing gap has increased in recent years. At the same time, by stimulating the impact of energy production and consumption, energy industry investment growth rate increased year by year. Overall, the interaction between the two is weakened, and the coordinated development is facing challenge.

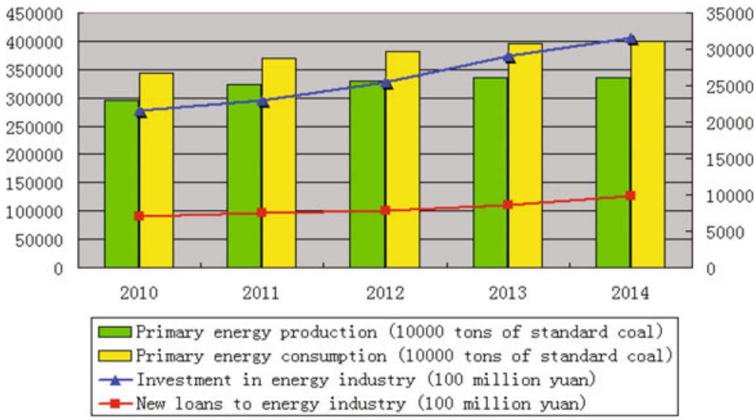


Fig. 1. Energy industry investment and financing, energy production and the evolution of energy consumption

The investments of oil and gas extraction, petrochemical, coking and nuclear fuel processing industry are unstable, lack of financing support. Overall, the energy industry internal financing coupling degree is not high.

3 The Coordination Model and Index System

According to the theory of synergetic, system always has a spontaneous irregular independence movement, while at the same time by the interaction of other subsystems, the connection between the seed system formed a cooperative movement. In the cooperative movement, there are many control parameters, which are divided into "fast variable" and "slow variable", that is, the order parameter is in the dominant position. With the constant change of the order parameter, the correlation between the subsystems is gradually increased before the critical point of the system; when the order parameter reaches the "threshold", plays a leading role in the association between subsystems system from disorder to order, from the chaos produced one kind of order structure. It can be said that the order parameter is the most prominent sign of the qualitative leap before and after the phase transition, which is about the characteristics and rules of the whole system transformation, find the order parameter of system is equivalent to find the observation of interaction between composite system movement and each subsystem of the key. Under the background of the greenhouse gas emission reduction and environmental protection policy, the energy industry investment subsystem has a multi-dimensional meaning, mainly including the existing production project pollution sources of investment, the new construction project "three simultaneous" investment, resource reserves and climate environmental protection policy, and so on four order parameters. Energy industry financing system is subject to monetary policy, energy industry financing and energy enterprise financial situation and so on three sequence parameters of the integrated drive. As a result, the coupling relationship of the energy industry investment and financing system is the sum of the various nonlinear relationships of the interaction between the order parameters. In order to eliminate the correlation between the subsystems within the control parameters, in this paper, the extraction and control parameter based, determine the order parameter estimated weight and measurement subsystem of order parameter, and then the composite system coupling coordination degree was evaluated.

3.1 The Theoretical Model of Investment and Financing in the Energy Industry

The coupling degree is used to describe the interaction degree of the energy industry investment and financing system, and to judge the degree of coordination and the stage of the coupling effect. First, establish the efficacy function. x_{ij} are the first I index of the control parameter of J, and x_{ij} is the function value of the function after the standardization. α_{ij} and β_{ij} are the upper and lower limit of the critical point order parameter of the system stability. Thus, "investment in the energy industry, energy industry financing coupled system orderly efficacy coefficient x_{ij} can be expressed as follows:

$$x_{ij} = \begin{cases} (x_{ij} - \beta_{ij})/(\alpha_{ij} - \beta_{ij}) \\ (\alpha_{ij} - x_{ij})/(\alpha_{ij} - \beta_{ij}), \end{cases} \tag{1}$$

where x_{ij} is the variable x_{ij} contribution to the system's function, which reflects the degree of satisfaction of the control parameters to achieve the goal, its range of values for [0, 1], 0 for the most dissatisfied, 1 for the most satisfactory. If $M(j = 1, 2, m)$ order parameter is extracted from the H control parameters in a subsystem $U(k = 1, 2)$, the order parameter matrix $(F_{ij})_{n\times m}$ with $n(i = 1, 2, \cdots, n)$ is formed. Because the investment in energy industry and energy industry financing are two different but interacting subsystems within the system, to achieve each order parameter of the order degree by the integrated method, the model is as follows:

$$u_k = \sum_{j=1}^{m} \theta_j \times F_{ij}, \quad \sum_{j=1}^{m} \theta_j = 1, \tag{2}$$

where θ_j are order parameter weights. Using the concept of capacity coupling and coefficient model for reference, the energy source is obtained, industrial investment and financing coupling degree function is:

$$C = \{(U_1 \times U_2)/\left[(U_1 + U_2) \times (U_1 + U_2)\right]\}^{\frac{1}{2}}, \tag{3}$$

where the letter C is the system interaction coupling degree, $C \in (0, 1)$; U_1 and U_2 represent energy industrial investment sub-system and the energy industry finance system on the system total contribution, namely investment in the energy industry comprehensive sequence parameters and energy industry financing order parameters.

For the energy investment and financing system, the significance of the coupling degree model is: quantitative description the interaction of between the two subsystems; reflect the order parameter of each subsystem in a certain period of time, the relationship between the number of regional interaction and adjustment process, in order to provide a basis for evaluation of the evolution trend of the interactive coupling of the composite system.

3.2 The Coordination Model of Investment and Financing in the Energy Industry

Coordination of each subsystem in the evolution process and factors of each part of the composition of different germplasm (order parameter), in which a unified whole showed mutual cooperation and harmony of property. If the order parameter of each subsystem is coordinated, it can promote the coordinated development of the whole complex system. In this paper, the difference of each subsystem in the composition of the composite system, the harmonious and consistent attribute is defined as the energy industry investment and financing Association scheduling. With respect to the coupling degree model, the coordination degree model considers the factors that influence the development speed of the overall order parameter of each subsystem, which makes up the coupling degree model only focusing on the horizontal coupling, and ignores the problem of the speed of development. It can be more comprehensive and accurate to judge the degree of coupling coordination between regional energy industry investment and financing system:

$$\begin{cases} D = \sqrt{C \times T} \\ T = \sqrt{\rho_1 U_1 \times \rho_2 U_2}, \end{cases} \tag{4}$$

where D is the coordination degree, $D \in (0, 1)$; C is the coupling degree; T is a comprehensive coordination index of energy industry investment and financing system, reflect the overall energy industry investment and financing synergy; ρ_1 and ρ_2 are undetermined coefficient; U_1 and U_2 are respectively the comprehensive order parameter of the energy industry investment and the order parameter of the energy industry financing. Coordination degree can be divided into 3 sections and 8 stages. As shown in Table 1 shows, through the comparison of the investment and financing system comprehensive order parameter, the sample can be roughly divided into three types: energy industrial investment excessive type, energy industry financing advanced type and investment and financing balanced development type, the following research will accordingly be identified and evaluated.

Table 1. The coupling coordination system and criteria of energy industry investment and financing

<table>
<tr><td colspan="3">$0 \leq C < 0.5$
Low level coupling interval
$0 \leq D < 0.5$
Offset recession interval</td><td colspan="3">$0.5 \leq C < 0.75$
Moderate coupling interval
$0.5 \leq D < 0.75$
Transition harmonic interval</td><td colspan="2">$0.75 \leq C < 1$
High level coupling interval
$0.75 \leq C < 1$
Coordinated development interval</td></tr>
<tr><td colspan="2">$0 \leq C < 0.3$
Low level coupling stage</td><td>$0.3 \leq C < 0.5$
Equally matched stage</td><td colspan="2">$0.5 \leq C < 0.6$
Preliminary running in stage</td><td>$0.6 \leq C < 0.75$
Depth running in stage</td><td colspan="2">$0.75 \leq C < 1$
Stage of benign coupling resonance</td></tr>
<tr><td>$0 \leq C < 0.1$
Extreme disorder decline stage</td><td>$0.1 \leq C < 0.3$
Moderate disorder decline stage</td><td>$0.3 \leq C < 0.5$
On the verge of a recession stage</td><td>$0.5 \leq C < 0.6$
Reluctantly coordination stage</td><td>$0.6 \leq C < 0.7$
Primary coordination stage</td><td>$0.7 \leq C < 0.75$
Intermediate coordination stage</td><td>$0.75 \leq C < 0.9$
Good coordination stage</td><td>$0.9 \leq C < 1$
Excellent coordination stage</td></tr>
<tr><td colspan="3">$U_1 > U_2$
Energy industry investment excessive type</td><td colspan="3">$U_1 = U_2$
The balanced development of investment and financing type</td><td colspan="2">$U_1 < U_2$
The energy industry financing in advance type</td></tr>
</table>

According to the connotation and characteristics of the investment and financing system of the energy industry, the comprehensive measurement index system is determined according to the principle of integrity, hierarchy and operation, see Table 2.

Table 2. Energy industry investment and financing coupling coordination system index system

Subsystem	Control parameter	Order parameter	Comprehensive order parameter
Energy industry investment subsystem	Total investment in environmental pollution control (100 million yuan)	Existing production project pollution source control investment	Comprehensive order parameter of investment subsystem
	Actual implementation of the "three simultaneous" project number		
	The actual implementation of the "three simultaneous" project investment (100 million yuan)		
	The actual implementation of the "three simultaneous" project environmental protection investment (100 million yuan)		
	Investment in energy industry (100 million yuan)	New construction project "three simultaneous" investment	
	Source of industrial pollution control projects completed investment in the year (100 million yuan)		
	Basic reserves of energy (100 million tons of standard coal)	Resource reserve	
Financing subsystem of energy industry	Carbon intensity of energy industry	Regulation of climate environmental protection policy	
	Total energy industrial assets (100 million yuan)	Financial situation of energy enterprises	Comprehensive order parameter of financing subsystem
	Main business income of energy industry (100 million yuan)		
	Total profit of energy industry (100 million yuan)		
	Energy industry should pay value-added tax (100 million yuan)		
	RMB loan balance (100 million yuan)	Monetary policy	
	RMB loans added (100 million yuan)		
	Balance of energy industry loans (100 million yuan)	Financing of energy industry	

4 Summary

Based on the system theory and synergetic theory, this paper establishes a coupled coordination model based on low carbon sustainable development and greenhouse gas emission reduction to analyze their interactions. The main conclusions are obtained as follows: (1) The investment and financing of the energy industry has a stress effect on the climate environment, and the climate environment has a binding effect on the investment and financing of the energy industry. The energy industry investment and financing stress effects on climate and environment are mainly: the energy industry investment and financing of energy production capacity increase the emissions of air pollutants; the constraint effect of the climate environment on the investment and financing of the energy industry is mainly caused by the severe climate and environmental protection policy. (2) Under the conditions of the policy of reducing greenhouse gas emissions, the interaction between the energy investment and financing has weakened year by year. In considering the greenhouse gas reduction factors, energy industrial greenhouse gas emissions in areas of high energy investment and financing coupling is relatively low, relatively poor coordination; Energy development and construction may not be an important cause of regional complex atmospheric pollution; The climate and environment protection policy constraints on energy investment is not strong, but has an important impact on energy financing.

Acknowledgement. This research is supported by the Program of National Natural Science Foundation of China under Grant [No. 71503216] and [No. 71401136], the Social Science Planning Project of Sichuan Province under Grant [No. SC15C051], the Central Universities Project under Grant [No. 2016NZYQN13], the Technology Finance and Entrepreneurship Financial Research Center Project under Grant [No. JR201507], and the System Science and Enterprise Development Research Center Project under Grant [No. xq15c02].

References

1. Du J, Bo Y, Yao X (2012) Selection of leading industries for coal resource cities based on coupling coordination of industry's technological innovation. Int J Min Sci Technol 22(3):317–321
2. Gneralp B, Seto KC (2008) Environmental impacts of urban growth from an integrated dynamic perspective: A case study of shenzhen, south china. Global Environ Change 18(4):720–735
3. Li Y, Li Y et al (2012) Investigation of a coupling model of coordination between urbanization and the environment. J Environ Manage 98(1):127
4. Norgaard RB (1990) Economic indicators of resource scarcity: a critical essay. J Environ Econ Manage 19(1):19–25
5. Smale R, Hartley M et al (2006) The impact of CO2 emissions trading on firm profits and market prices. Climate Policy 6(1):31–48
6. Tang Z (2015) An integrated approach to evaluating the coupling coordination between tourism and the environment. Tourism Manage 46:11–19

7. Wang SJ, Ma H, Zhao YB (2014) Exploring the relationship between urbanization and the eco-environmentła case study of beijing-tianjin-hebei region. Ecological Indicators 45(5):171–183
8. Weick KE (1976) Educational organizations as loosely coupled systems. Administrative Sci Q 21(1):1–19

Space-Time Analysis for Water Utilization Efficiency of Urban Agglomeration and Its Economic Influence Factors łin Chengdu Urban Agglomeration

Yeyun Huang[1] and Yunqiang Liu[2(✉)]

[1] College of Economics, Sichuan Agricultural University, Chengdu 611130, People's Republic of China
[2] College of Management, Sichuan Agricultural University, Chengdu 611130, People's Republic of China
liuyunqiang@sicau.edu.cn

Abstract. This paper uses the DEA method to establish the evaluation model of water resources efficiency to make an empirical study on the urban agglomeration of Chengdu city in 2008–2014. The results show that the efficiency is increasing year by year, and the economic efficiency and environmental efficiency of the water resources utilization are improved. Meanwhile, under the control of the technology, education and transportation level to analyze of efficiency from the economic factors, this paper gets the conclusion that: (1) the city water use efficiency and the level of economic growth showed a significant linear correlation. (2) Technology shows no obvious linear relationship with the efficiency of water use. Traffic enhances the effect of openness on water efficiency. Education is negatively correlated to the efficiency. It indicates that the urban agglomeration should enhance the traffic infrastructure on its regional economic advantage, optimizing the industrial.

Keywords: DEA data envelopment · Water use efficiency · Urban agglomeration · Economic development · Econometric analysis

1 Introduction

With the economic development and population growth, water shortage is a severe problem. According to the National Bureau of statistics report, Sichuan water consumption for each industrial added value dropped from 135 m^3 per million yuan in 2008 to 46 m^3 per million yuan in 2014, a decline of 66%, achieving the requirements of the State Council in advance. However, compared with Beijing, Suzhou and other "national water-saving city" in recent years, which maintain the value of 16–19 m^3 per million yuan, the efficiency of water use needs improvement.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_102

The shortage of water resources is a global problem, and how to make efficient use of water to ease the supply and demand of resources is urgent. The coordinated development of economy and environment is the foundation of building a resource-saving society. On the one hand, the main branches of the economy have an impact on the efficiency of water use [2], on the other hand, the improvement of water use efficiency is conducive to the local food exports [7].

Therefore, the quantitative analysis of economic development with water use for sustainable development is particularly important. Previous studies had focused on the evaluation of the efficiency of water use on agriculture, industry and mixed model, or on the profound the model. The present studied the use of water in the economic output perspective, ignoring the ecological impact, this paper aims to evaluate the Chengdu city group water utilization through two aspects of economic and environmental efficiency. Few existing research attached quantitative analysis of factors affecting the efficiency of water use, this paper based on the empirical analysis study the efficiency of water use and its economic factors, prolonging sustainable development for the construction of water-saving.

2 Research Objectives

There are many ways to analyze the utilization of water resources, in view of the water potential of the unified understanding degree is not high. The range of fluctuation is difficult to quantitative analysis of the decoupling theory. This article based on the perspective of the utilization efficiency of water resources to study the water use in the city of Sichuan province. This article unfolds as follows. Primarily, In order to explore the relationship between water use efficiency and economic development, this article based on the Cobb-Douglas production function, to set the input-output model to measure the efficiency of water resources utilization. Secondly, according to the difference of water use in time and space, this paper divides the comprehensive utilization efficiency of water resources into economic efficiency and environmental efficiency. Thirdly, this article analyzes the impact of economic development on the efficiency of water use, to discover the relationship between economic development and water consumption in urban agglomeration. Lastly, In order to reflect the economic indicators in a better way, this paper under the control of the city of scientific technology, education and water transportation level, to do further analysis between economic level and efficiency of water use. It's of great importance to profoundly analyze economic development and resource allocation. Figure 1 shows the construction.

This paper chooses BBC model to analyze the relative efficiency of variable returns to scale. The BBC model increases the convexity assumptions on the basis of the CCR model, which will be helpful to analyze the environmental and economic output [4,8,9]. We add the curve measure to the BBC model. The curve measure is the environmental efficiency evaluation method of nonlinear analysis, analyze the output with radial infer, use reciprocal curve to measure the pollutants, in order to achieve efficiency evaluation in the process of increasing the output and reduce the pollution.

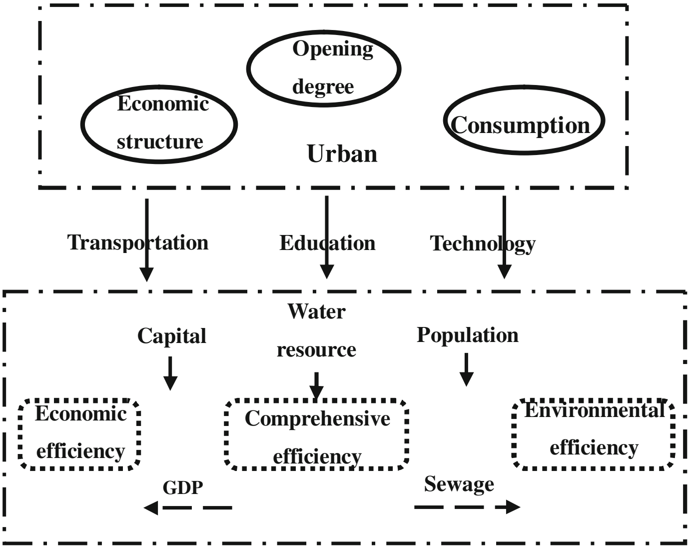


Fig. 1. The construction of the paper

In order to fully reflect the input and output of water resources in cities, the paper based on the Cobb-Douglas production function, constructed an input output analysis framework including water use and water pollutant emissions. The exact input-output index system followed in Table 1. The fixed assets investment of the whole society reflects the scale of capital investment of the whole society in a period. The local population, compared to the employment population, better expresses the contribution to the increase in water consumption and economic output. We use the annual water supply as the water consumption of each grade city. All the data comes from the Sichuan statistical yearbook. This paper adds city sewage emissions as undesirable outputs. In case of pollution and other undesirable products are produced, based on the previous [3,4] for pollutants treatment, this paper inserts the reciprocal evaluation which has comparative advantage to DEA model as the environmental impact.

Table 1. Input-output index system

Input	Water resource	Cities' water consumption
	Capital size	Social investment in fixed assets
	Population	Current population
Output	Economic output	Current GDP
	Environmental effects	Current sewage discharge

Table 2. The unit root test for panel data

	Structure	Export	Consumption
Statistic	−5.0904	−5.7898	−1.30E+02
Z	−4.5241	−3.7001	−4.9723
Prob.	0	0	0

3 Space-Time Analysis for Efficiency

The paper uses the DEAP 2.0 to set the model, lists the results in Table 3. It is obvious from Table 3 The water use efficiency of Chengdu urban agglomeration increased gradually in 2008 to 2014. The water utilization efficiency of Ziyang in 2008 to 2012 showed the optimal level 5 times, indicating that the capital investment, labor and water resources allocation in Ziyang city is in a good condition and take the best place among Chengdu urban agglomeration. We use ArcGIS to demonstrate the variation of the comprehensive efficiency.

Table 3. 2008–2014 Chengdu urban agglomeration water use comprehensive efficiency

	2008	2009	2010	2011	2012	2013	2014
Chengdu	0.76	0.69	0.81	0.87	0.9	0.92	1
Deyang	1	0.63	0.84	1	0.98	0.99	1
Mianyang	0.72	0.53	0.62	0.75	0.79	0.83	0.87
Leshan	0.81	0.73	0.76	0.9	0.89	0.86	0.9
Meishan	0.71	0.66	0.72	0.85	0.81	0.81	0.85
Yaan	0.65	0.73	0.64	0.8	1	1	1
Ziyang	1	0.86	0.67	1	1	1	1

On Fig. 2, these directly show the trend of water utilization efficiency. From the perspective of time series, the financial crisis in 2008 affected the economic development and water use of Chengdu urban agglomeration which led to the tendency of water utilization efficiency declined in 2009. In 2011, with economic recovery and city industrial construction, water utilization efficiency and the indirect influence gradually recovered to the level of 2008.

From the perspective of the spatial distribution of the industry, Chengdu urban agglomeration pillar industries (electronic information, water and electricity, pharmaceutical and chemical industry, machinery, metallurgy, food and beverage, etc.) of large enterprises were mainly distributed in Chengdu, Deyang, Mianyang, Ziyang, However, the development of industry in Meishan, Leshan lagged behind, which lacked the support for large enterprises and projects.

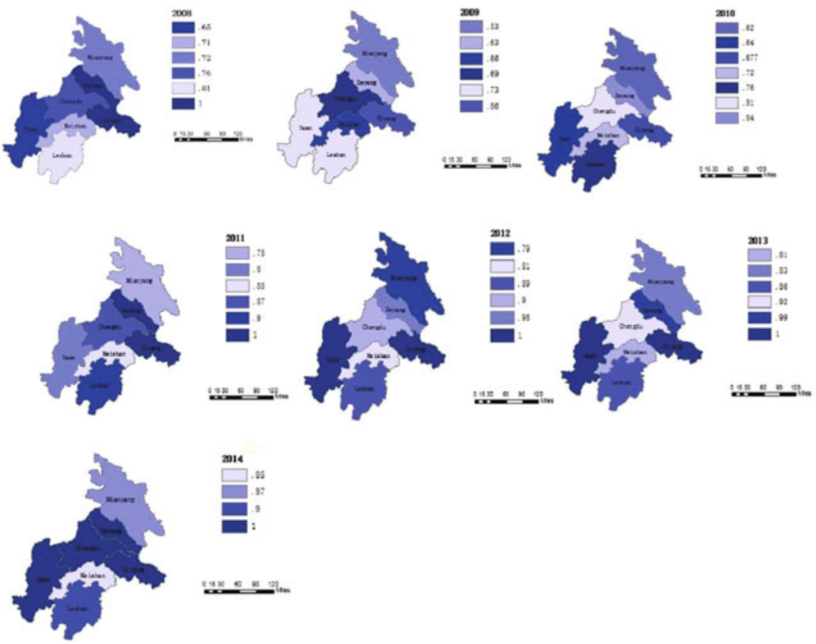


Fig. 2. Chengdu city group water use comprehensive efficiency from 2008–2014

This paper analyzes the comprehensive utilization efficiency of water resources from two aspects of economic efficiency and environmental efficiency. Environmental efficiency reflects the scale is insufficient or redundant. Based on the conventional expectations output, we adds the new sewage emissions as undesirable output. It not only reflects the moderation between water supply quantity and capital investment, but also reflects the reduction of potential pollutant emission under the same factor input and output conditions [6,10]. It is of great help for the various regions to establish future planning objectives in the existing economic and environmental conditions. Economic efficiency refers to the existing economic input, the best use of technology or the ability to deal with pollutants. It embodies the ability to achieve the maximum output under the conditions of a variety of input factors. In this paper the DEA-BBC input-output model, the economic efficiency reflects the investment of fixed assets and water resources social investment which leads to the ability to achieve the maximum economic output in the certain number consumption of a city. Analysis of the economic efficiency and environmental efficiency of water resources in Chengdu urban agglomeration by DEAP 2 is shown in Fig. 3. Based on the time series, it is obvious that 2009 was a turning point due to the financial crisis in 2008. In terms of the city's water efficiency, Yaan's economic efficiency had been maintained at a high level. The reason is that the economic development of Yaan lagged behind by the state in the Chengdu urban agglomeration. Its economic output and degree of industrialization were low, thus embodying water utilization efficiency good situation in the input and output. Ziyang and Deyang kept a stable efficiency at the high level in 2011. Deyang's pillar industry was mechanical, which accounted for nearly half share. Although the development of machinery industry was still in the extensive stage, under the atmosphere of 2011–2014, tap estates had a greater contribution to pull the local economic output. Ziyang on the location was the unique regional center city which connected Chengdu and Chongqing the "dual core". It's known as its automobile industry the same as Deyang, which drove the rapid economic development through increasing industrial output then indirectly improved the economic efficiency. Compared to Mianyang, the city manifested in industrialization as well, it existed an obvious gap. Mianyang played a relatively low ecoefficient role among Chengdu urban agglomeration with its electronic as pillar industry. The falling behind velocity of the increasing development restricted its economic efficiency to a higher level. Although Mianyang regarded as the second biggest city in Sichuan, the lowest water environmental effect dragged it at the bottom and pull down its water utilization comprehensive efficiency. Similarly, Chengdu, which was specialized in electronics, medicine and so on, because it was the urban agglomeration center city and the capital city of Sichuan Province, its economic development was much higher than that of other prefecture level cities which raised its water use efficiency. In terms of the environmental efficiency of water use, Chengdu urban agglomeration showed the effective in environmental efficiency but relatively low valid in eco-efficiency. The reason is that it didn't make the best use of factors. For illustration, Yaan and Meishan did not serve its turn. Meanwhile, the plenty of cities had already taken a big scale, such as Chengdu, Deyang, Leshan, Ziyang and Mianyang. In general, the input-output ratio was reasonable, and the combination of factors achieved a certain scale economy. Therefore, it is close to the optimal efficiency on the scale of the city but insufficient in eco-efficiency.

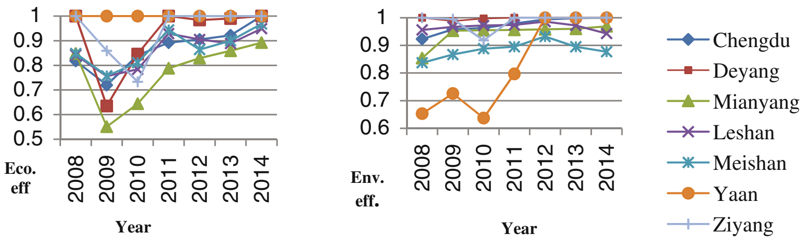


Fig. 3. 2008–2014 Chengdu urban agglomeration water use Economic and Environmental efficiency

4 Economic Effects

4.1 Hypothesis

Based on the above empirical, it is visible that economic output and its level of city development has close relationship with water utilization.

This paper assumes that the explanatory variables are: economic structure, opening degree, consumption level. At the same time, this paper selects the technology, transportation and education level as the control variables.

Hypothesis 1: economic structure has a negative impact on water use efficiency. The economic structure is the composition and structure of the national economy. The influence of economic structure on water resources utilization efficiency can be analyzed from two aspects. On the one hand, the industrial production have positive effects on the economy of a region, the economic structure in the industry tend to have higher economic output which shows high utilization efficiency of water. On the other hand, the increase in industrial investment will increase the impact of the ecological environment, resulting in a larger proportion of environmental output, thereby reducing the city's water use efficiency. Now the negative effect of the development of Chengdu city group is the ecological environment influence over its economic effects is unknown.

Hypothesis 2: the degree of opening to the outside world has a positive impact on the efficiency of water resources utilization.
The degree of opening effects the import of resources and the introduction of new technologies. Chengdu is located in China's western region, the western energy technology itself is relatively backward. The change in the degree of dependence on foreign trade has great influence on the west. Therefore, the study on the degree of opening to the outside world is of great significance to the utilization efficiency of water resources in Chengdu urban agglomeration.

Hypothesis 3: consumption level has a positive impact on water use efficiency.
To some extent, the consumption level reflects the scale of production and indirectly affects the efficiency of resource utilization. On the one hand, consumption is an important booster of economic growth. On the other hand, all kinds of garbage generated in the process of consumption also makes the decline in environmental quality [1]. Whether the ecological impact of rising consumption level is beyond its economic effects? This paper assumes that the consumption level is a positive impact on water use efficiency.

4.2 Analysis of Indicators

This paper will explain the following three aspects on the selection of economic factors and the setting of variables: As explained variable, utilization represents the regional comprehensive efficiency of water use. The data is calculated by the DEA-BBC model. Other data comes from the 2009–2015 Sichuan statistical yearbook. For Explanatory variables, structure refers to the economic structures, this paper uses the sum of the primary and secondary industries with the ratio of the service industries to express the economic structure. Set the ratio due to the water consumption of primary and the secondary industriesshare more, the service industrial (including commodity trade, catering and accommodation, transportation organizations and other services industry) accounted for only 3% (Sichuan water resources bulletin 2014). In terms of the gross national product structure, in the Chengdu urban agglomeration, the primary industry occupies less compared to the secondary, the service industries. Therefore, use the sum of primary and secondary industrial added value divided by the added value of the third industry to represent the economic structure and investment allocation direction. Consumption indicates the level of consumption, this paper selects the total retail sales of social consumer goods to express the transaction situation of various organs, enterprises and institutions, in order to reflect the domestic consumption level and demand of the city. Export is used to indicate the degree of opening up. In this paper, the ratio of total import and export to GDP is used to reflect the degree of opening up. The city's total import and export of goods better embodies the foreign trade level. Because of its joint venture, said the unit is million, taking into account other indicators to billions of units, the direct use of the total import and export data analysis model will have a certain deviation, and the change of exchange rate is floating, it is difficult to calculate by import and export the amount of US dollars into RMB. This paper selects the import and export volume and the current ratio of GDP to reflect the effect of the level of international trade.

Table 4. Definition and description of variables

Variables	Definition	Description	Effect
Utilization	Water utilization efficiency	Local comprehensive efficiency of water use	−
Structure	Economic structure	(Primary + Secondary)/(Service)	+/−
Consumption	Level of consumption	Total retail sales of social consumer goods	+/−
Export	Degree of opening up	(Import + Export)/(Current GDP)	+
Technology	Level of technology	The ratio of technology expenditure	+
Education	Level of education	The ratio of education expenditure	+
Transportation	Level of transportation	The ratio of transportation expenditure	+

For the control variables, this paper use the proportion of public finance expenditure to reflect the levels. The city's technology investment has positive effect on improving the efficiency of water resources utilization. The progress of technology can create more water-saving facilities, increase the circulating water, thereby increasing the utilization efficiency of water resources. The educational level of the city reflects a city's progress and the civilized degree. The level of urban transportation will affect the degree of opening up. At the same time, the improvement of the level of transport will also play a role in alleviating the imbalance of water resources distribution [5], so as to improve the utilization efficiency of water resources (Table 4).

4.3 Model Description

The econometric model of panel data includes fixed effect model, random effect model and least square regression method. The equitation is as followed.

$$\text{Utilization} = b_1 + b_2 \times \text{Structure} + b_3 \times \text{Export} + b_4 \times \text{Consumption} + \varepsilon, \quad (1)$$

where b_1 stands for the intercept, b_2, b_3, b_4 represent the coefficient of the economic structure, the degree of opening up and the social consumption. ε is the stochastic error. Because of the regional difference of the urban agglomeration is relatively small, the difference of the water resources utilization efficiency between different sections of the cross section can be ignored. The time span is small, and the variation of parameters can be ignored as well. In this paper, WLS regression is used to fit and analyze the influencing factors and weights, and the control variables are introduced gradually.

$$\begin{aligned} \text{Utilization} &= b_1 + b_2 \times \text{Structure} + b_3 \times \text{Export} + b_4 \times \text{Consumption} \\ &\quad + b_5 \times \text{Technology} + \varepsilon, \end{aligned} \quad (2)$$

$$\begin{aligned} \text{Utilization} &= b_1 + b_2 \times \text{Structure} + b_3 \times \text{Export} + b_4 \times \text{Consumption} \\ &\quad + b_5 \times \text{Technology} + b_6 \times \text{Transportation} + \varepsilon, \end{aligned} \quad (3)$$

$$\begin{aligned} \text{Utilization} &= b_1 + b_2 \times \text{Structure} + b_3 \times \text{Export} + b_4 \times \text{Consumption} \\ &\quad + b_5 \times \text{Technology} + b_6 \times \text{Transportation} \\ &\quad + b_7 \times \text{Education} + \varepsilon, \end{aligned} \quad (4)$$

where b_5, b_6, b_7 represent the effective coefficients of the technology, transportation and education towards the water use efficiency.

4.4 Empirical Analysis

(1) Test of stationary

In order to avoid spurious regression, and ensure the validity of the estimation results, we must check the stability of each panel sequence. We use Stata 11.0 to perform unit root test for short panel data. The sample capacity is small, the effect of the unit root test for single variable is weak, thus we use the LLC test to analyze the data. The original hypothesis H_0 has the same unit root. The test results are shown in Table 2, the data has a high degree of significance, so the result is to reject the original hypothesis. The perturbation terms of different individuals are independent of each other, and the data are stable.

(2) Results and analysis

According to the significance of the data, the mixed OLS model is better than the fixed effect and random effect by Stata 11, and the Table 5 is the regression fitting result of the three models. It was the primary and the secondary industry high proportion improve the utilization efficiency of water use. In addition, the further openness would introduce new energy, advanced technology and talents, and it could largely reduce economic output and has high impact on the ecological environment. The regression coefficient of the level of consumption is significantly positive. We use stepwise least squares regression, in the case of abandoning the constant, to analyze the standardized coefficient.

Table 5. Comparison of regression coefficients

	Before the insert of control variables				After the insert of control variables		
	Pooled OLS	Fixed effect	Random effects	WLS	+ Technology	+ Transportation	+ Education
b_2	0.217[a]	0.230[a]	0.2259[a]	1.041[a]	0.404	0.379	0.700[a]
b_3	4.162[b]	4.421	4.3012[b]	0.532[b]	0.671[a]	0.939[a]	1.253[a]
b_4	0.001[a]	0.001[b]	0.001[a]	0.637[b]	0.981[b]	0.607[a]	0.542[a]
b_5	–	–	–	–	−0.1	–	–
b_6	–	–	–	–	–	0.418[a]	0.471[a]
b_7	–	–	–	–	–	–	−0.338[a]
R^2	0.412	0.411	0.411	0.407	0.354	0.503	0.577
F	55.31	6.61	–	10.759	6.995	12.895	13.649

[a] Correlation is significant at the 0.01 level, [b] Correlation is significant at the 0.05 level.

From Table 5, it is obvious that the economic structure takes the greatest place among the economic indicators to demonstrate the trend of the water use efficiency. The reason is that the Chengdu urban agglomeration has a higher degree of industrialization. The domestic consumption promotes economic diversification of functions of the city through the sharing of resources among cities, which indirectly improve the utilization of water. Meanwhile, international trade will bring the inflow of technology and talent, and promote the development of economy.

In this paper, the coefficient of technology is negative as control, and it didn't pass the significance test. This doesn't mean that there is no linear relationship between the technology and the efficiency of water use. It is based on the technology proportion of public expenditure. In fact, many of the city's scientific development depend not only on the public finance expenditure, it depends on technological innovation in enterprises and large companies. The traffic level shows that increasing its proportion of public finance expenditure has a positive effect on improving the efficiency of water use. Under the influence of the traffic, the effect of the economic structure weakened, and the effects of opening up rise up. The coefficient of educational level is negative.

(3) Heteroscedastic test

In order to check the deviation between the fitting value and the real value of the model and to ensure the regression parameter estimation has good statistical properties, the paper makes use of the residual graph to illustrate the heteroscedasticity test. From the residual plot of heteroscedasticity test that, in general, the fitting degree of the model is good. The fit of economic structure is good, the scattered distribution is random and uniform (Fig. 4).

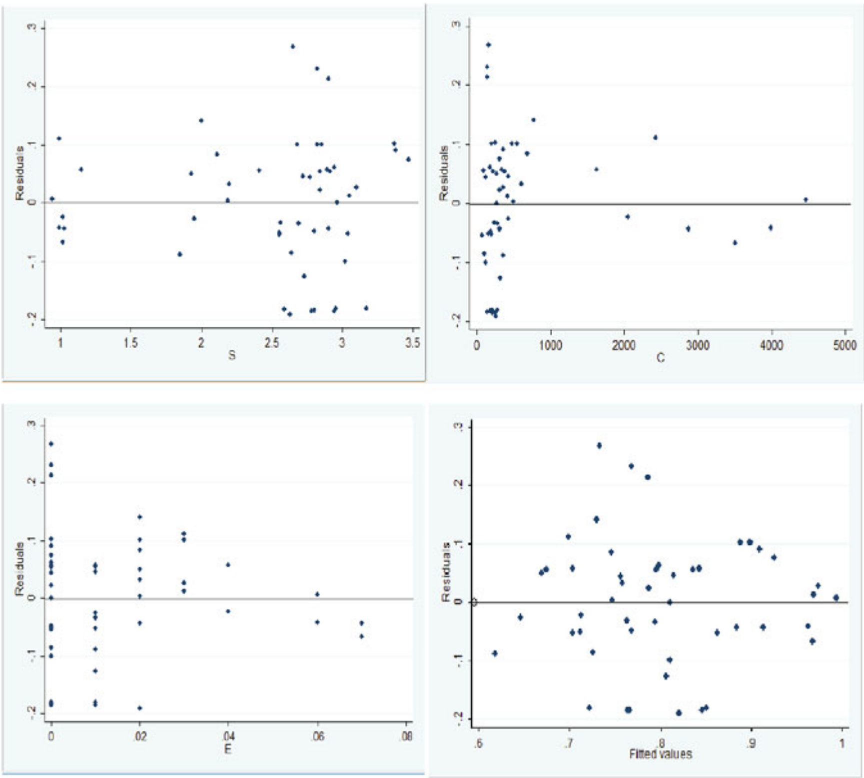


Fig. 4. The residue of structure, consumption, openness and fitted values

5 Conclusion

The water use efficiency of Chengdu urban agglomeration has difference in time and space. Attach the time series as the axis, since the financial crisis into the economic buffer period, the efficiency of water use was gradually increasing. From the perspective of spatial difference, because of its geographical advantages, in the electric, automotive and other industries as a pillar industry in the city, its water use efficiency was better. For some cities, although the marginal scale investment was higher, due to its serious environmental pollution and relatively weak economic growth, it shared the low utilization efficiency of water use. The data shows that the water use of Chengdu urban agglomerations still had some room to improve, the government departments should make efforts to increase sewage treatment.

The water utilization efficiency from the perspective of economic level analysis, found the water resource utilization and economic development is linear, the economic structure takes the most profound effect. Increasing the utilization rate of resources depends largely on the green industry output; the opening degree and city's consumption affected the input of technology and talent to indirectly affect the efficiency of water use. Meanwhile it directly affect the water use through the flow of resources among the agglomeration.

This paper makes a further analysis on the urban development, and finds that the level of urban transportation has a positive effect on the efficiency of water resources utilization and will increase the effective proportion of the degree of opening up. Raising the level of education in the short term is not conducive to the water use. Due to the long period of payback in the educational investment, it requires for a higher levels of openness to compensate the effect for extra education spending of the utilization efficiency of water resources.

Acknowledgements. This study is supported by Ministry of Education Humanities and Social Sciences Youth Project (15YJC630081); Sichuan Provincial Social Science Research "12th Five-Year" planning project (SC14C027); Sichuan Provincial Institute of systems science and Enterprise Development Research Center (Xq14B08).

References

1. Kox HLM (2013) Export decisions of services firms between agglomeration effects and market-entry costs. Service industries and regions. Springer, Heidelberg, pp 177–201
2. Kulmatov R (2014) Problems of sustainable use and management of water and land resources in Uzbekistan. J Water Resour Prot 6(1):35–42
3. Liao H (2011) Utilization efficiency of water resources in 12 western provinces of China based on the DEA and malmquist TFP index. Resour Sci 33(2):273–279
4. Mai Y, Sun F et al (2014) Evaluation of China's industrial water efficiency based on dea model. J Arid Land Resour Environ 11:008
5. Malaeb L (2004) Constituent transport in water distribution networks
6. Melo PC, Graham DJ, Noland RB (2009) A meta-analysis of estimates of urban agglomeration economies. Reg Sci Urban Econ 39(3):332–342
7. Taheripour F, Hertel TW et al (2016) Economic and land use impacts of improving water use efficiency in irrigation in South Asia. J Environ Prot 7(11):1571–1591
8. Thanassoulis E (2000) Dea and its use in the regulation of water companies. Eur J Oper Res 127(1):1–13
9. Veettil PC, Ashok A et al (2011) Sub-vector efficiency analysis in chance constrained stochastic dea: an application to irrigation water use in the Krishna River Basin, India. In: The 122nd Eaae Seminar "Evidence-Based Agricultural And Rural Policy Making: Methodological And Empirical Challenges of Policy Evaluation"
10. Wu F, Liu Z (2008) Research on the mechanism of how city group drive economic growth-empirical evidences from 16 cities of Yangtze River delta. Econ Res J 11:126–136

Exploring Linkages Between Lean and Green Supply Chain and the Industry 4.0

Susana Duarte(✉) and Virgilio Cruz-Machado

UNIDEMI, Departamento de Engenharia Mecânica e Industrial,
Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
scd@fct.unl.pt

Abstract. A new paradigm with a considerable influence on the industrial value creation is begun to evolve. Industry 4.0 represents the development to the fourth industrial revolution. A new shift in business and technology trends is confirmed. With this new paradigm of high-tech industry, supply chains will develop into highly adaptive networks. Lean and Green supply chain management should have better opportunities to be performed, becoming more efficient and competitive. The aim of this study is to link the lean and green supply chain characteristics to the Industry 4.0. This paper has twofold: first, presents an overview of the phenomenon Industry 4.0 and the lean and green supply chain management. Next, a conceptual model is developed which incorporates the Industry 4.0 topics to the well-known lean and green supply chain. This paper provides an understanding of the role of lean and green paradigms in the new era of industrialization.

Keywords: Industry 4.0 · Lean paradigm · Green paradigm · Supply chain

1 Introduction

Today we face a new paradigm introduced by Germany in 2011, the Industry 4.0. This represents the beginning of the fourth industrial revolution and is driven by modern information and communication technology (ICT) [21,26,27]. By Industry 4.0 intends the optimization of value chains by implementing an autonomously controlled and dynamic production [21], through a full automation and digitalization processes. This paradigm is based on the idea that communication via internet allows a continuous exchange of information applying cyber-physical systems (CPS) [27,30]. The CPS provides the source for the creation of internet of things and services and their combination makes possible the Industry 4.0 [15,27]. CPS integrates networking, computation and physical processes and take them together to create a global value chain networked [15,27,30].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_103

According to [34] to establish the global value chain networks, the Industry 4.0 describes a production oriented CPS that integrates production facilities, warehousing and logistics systems and even social requirements. In addition, Germany Trade & Invest (GTAI) [15] mention that the industrial value chain, product life cycles and business information technology combination must integrate the processes from the product design to production, supply chain management, aftermarket service and training [15]. An intelligent factory is in development and is coined as smart factory. With this concept, others appear and are important for the establishment of implementation of Industry 4.0, as for example smart products, smart manufacturing and smart data.

The authors [27] mentioned that this implementation is still in progress. That's why it is important to understand the role of lean and green supply chain management in Industry 4.0. For example, the lean waste would be recognized with the smart factory implementation [28]. The resource efficiency which is a lean and green concept, are in the focus of the design of smart factories [27]. This study intends to understand if the Industry 4.0 allows the lean and green supply chains concepts become more important. That is, if it enables more easily the deployment of lean and green characteristics. A number of characteristics were presented on model, namely: (i) manufacturing, (ii) logistics and supply, (iii) product and process design, (iv) product, (v) customer, (vi) supplier, (vii) employee, (viii) information sharing and (ix) energy.

The remainder of this paper is organized as follows: in Sect. 2, a theoretical background on Industry 4.0 and lean and green supply chain are presented; in Sect. 3, a combination between lean and green supply chain and Industry 4.0 is developed; Finally, some concluding remarks are drawn.

2 Industry 4.0 Paradigm

The Industry 4.0 is considered the paradigm of the fourth stage of industrialization and describes a vision of future production [21,30]. The core idea of Industry 4.0 is the integration and application of information and communication technologies to implement Internet of Things and Services so that business process and engineering process are deeply integrated making an environment intelligent [28,34]. The concept of industry 4.0 which represents the integration of the virtual and physical worlds in a way that together create a truly networked environment and where intelligent objects communicate and interact with each other, is a Cyber-physical systems [15]. According to [19] the Industry 4.0 "will involve the technical integration of CPS into manufacturing and logistics and the use of the Internet of Things and Services in industrial processes. This will have implications for value creation, business models, downstream services and work organization."

The Industry 4.0 is represented by three features [19,30,34]: (i) horizontal integration across the entire value networks; (ii) vertical integration and networked manufacturing system; and (iii) end-to-end digital integration of engineering across the entire value chain or product life cycle.

The horizontal integration across the entire value network refers to the integration of the various systems used in the different stages of the manufacturing and business planning processes that involve an exchange of materials, energy and information both within a company as logistics, production, and marketing, and between several different companies; The idea is that information, material and money can flow easily among different companies creating new value networks as well as business models. This can result in an efficient ecosystem [19,30,34].

The vertical integration refers to the integration of the various information and physical systems at the different hierarchical levels, as for example the production management, manufacturing and execution, and corporate planning. This integration is inside a factory to create flexible and reconfigurable manufacturing system [19,30,34].

The goal of the horizontal and vertical integration is to deliver an end-to-end solution. The end-to-end solution refers to the digital integration of engineering across the entire value chain to support product customization: from the raw material acquisition to manufacturing product, and product in use and in the end of life [19,30,34].

Through these features, the Industry 4.0 expects to implement an environment more flexible, efficient, and sustainable. The idea is to individualize the customer requirements, as a customized product through a mass customization, improving productivity and achieving higher levels of quality with a manufactured profitably result [6,19]. Indeed, by applying advanced information and communication technologies and systems in the manufacturing and supply chain operations, the industry 4.0 addresses the smart factory [28]. Smart factory is designed according to sustainable and business practices, insisting upon, flexibility, adaptability and self-adaptability, learning characteristics, fault tolerance, and risk management [15]. Therefore, standards are essential to ensure the exchange of data between machines, systems and software and guarantee that product moves within a network value chain [6].

That is, high levels of automation come as standard [15]. Automation systems, manufacturing and product management are integrated and are the base of the smart factory [6]. Manufacturers can now add sensors and microchips to tools, materials, machines, vehicles and buildings to communicate with each other in real-time to make smart products [15].

According to [36], "products know their histories and their routes, and thereby not only greatly simplify the logistic chain but also form the basis for product life cycle data memories". Also, the products can be manufactured because smart factory is being supplied with energy from smart grids [30].

Not only smart factory and smart product were defined in this new industrialized era. Others concepts connected to them are considered in the literature. For their work development, Kolberg and Zuhlke [21] considered four different smart concepts to define the smart factory, namely, smart planned, smart product and smart machine, and smart operators. The authors [30] make mention of the smart grid, smart logistics and smart data. Sanders et al. [28] mention

Table 1. Concepts of Industry 4.0

Concepts	Description
Smart factory	Smart factory represents the key characteristic of Industry 4.0 [15]. The smart factory will be more flexible, dynamic and intelligent [27], where people, systems and objects communicate with each other [15]. The internet of things and services are the main enabler technology for a smart factory [15, 29]
Smart manufacturing	Manufacturing will be equipped with sensors and autonomous systems which allow that operations can be optimized with a minimum employee's intervention [27, 29]. It produces small-lot products of different types, more efficiently [34]
Smart product	A smart product is a product with sensors and microchips that allow communication via the internet of things with each other and with employees [27]. It holds the information about its requirements for the manufacturing processes and manufacturing machines [21, 30]
Smart logistics	It is one of sustainable mobility strategies [15]. Smart logistics will use CPS for carrying the material flow within the factory and in the supply chain (between factories, customers and other stakeholders) [30]. The transport equipment is a part of smart logistics that is able to react to unexpected and autonomously should be able to drive between the starting point and the destination [30]. Distribution and procurement will increasingly be individualized [27]
Smart engineering	Includes product design and development, production planning and engineering, production and after sales service [29]
Smart data	Smart data is structured information of data that can be used for decision-making [30]
Smart machine	Machines and equipment will have the ability to improve processes through an intelligent decision-making, instead of being directly instructed [27, 34]. The smart machines should have additional autonomy and sociality capabilities to adapt and reconfigure to different types of products [34]
Smart planner	Smart Planner optimizes processes in real-time [21]. Decentralized self-organization [27]
Smart operator	smart operator is an employee who supported by ICT, control and supervise ongoing activities [21]. Employees can be quickly directed to the right tool [6]
Smart customer	Customers' needs and behaviors are analyzed in real-time in way to provide them with new and more sustainable products and services [29]
Smart supplier	Based on factory needs it is possible to select the best supplier (which allows higher flexibility) and strengthen a sustainable relations with suppliers (by increase information sharing in real-time) [29]
Smart grid	Responsible to supply energy to a factory [30]. Energy management [15]
Smart energy	Monitor and provide feed-back on energy production and use [23]

others concepts as the smart systems, smart environment, smart machine and smart devices, and smart task. Table 1 compiles several concepts of Industry 4.0.

Through the integration of the industry concepts and technologies it should be possible provide a customized or individualized product or service and at the same time be highly adaptive to demand changes [15]. These changes must be made on all stages of product life cycle: design phase, raw material acquisition phase, manufacturing phase, logistics and supply phase, and the use and end of life phases [15,30]. Therefore, the requirements for design and operations of our factories become crucial for the success [36].

3 Lean and Green Supply Chain Paradigms

Nowadays, lean and green supply chain is an integrated approach; they have different objectives and principles but they complement each other [7,9,10,12–14,31]. Lean supply chain is about to increase value for customers by adding product or service features, with the elimination of waste or non-value steps along the value chain [11]. Green supply chain regards to reducing environmental impacts and risks while improve ecological efficiency of the organizations and their partners, and try to achieve corporate profit and market share objectives [35].

These two paradigms are often seen as compatible because of their joint focus on waste reduction [5]. Lean paradigm is concerning to the elimination of waste in every area of design, manufacturing, and supplier network and factory management [13]. The basic forms on the reduction and elimination of waste are [17]: production, waiting, transportation, unnecessary inventory, inappropriate processing, defects and unnecessary motions. One more waste is pointed by [31] as the unused employee creativity. Green considers ways to eliminate waste from the environment's perspective [11]. The waste generation have the form of [16]: Greenhouse gases, eutrophication, excessive resource usage, excessive water usage, excessive power usage, pollution, rubbish and poor health and safety. In their research [12] mention that the two paradigms have the same type of wastes: (i) inventory; (ii) transportation, and (iii) the production of by-product or non-product output. According to [5] the removal of non-value adding activities suggested by lean paradigm can provide substantial energy savings which integrates the principles of green paradigm.

The combination of lean and green supply chain practices have better results than the total from the implementation of each, but separately [12]. The two paradigms have similar characteristics. According to [4] both paradigm practice contribute for: (i) the increase of information frequency, (ii) the increase of the level of integration in supply chain, (iii) the decrease of production and transportation lead time, (iv) the reduction in the supply chain capacity buffers; (v) and the decrease of inventory levels. Another practice that contributes for the better employ and use of all tools is the involvement of the employees [12]. Both paradigms look into how to integrate product and process redesign in order to prolong product use, to allow easily the recycling or re-use of products, and to make processes with less wasteful [12]. In the supply chain both paradigms ask for a closed collaboration with partners [22]. In addition, waste reduction, lead time reduction, and use of techniques and approaches to manage people, organizations, and supply chain relations are synergies mention by [13].

Commitments must be made within factory, supplier network and customer, for the better deployment of lean and green practices in way to achieve the best supply chain efficiency. In the authors previous study [11] it was presented a table

Table 2. Lean and green paradigms characteristics in supply chain

Characteristics	Lean	Green
Philosophy focus	Long-term thinking [2]	Long-term thinking since environmental impacts affects the natural environment for many years [18]
Market focus	Serve only the current market segments with predictable demand [33]	Demands from at least some customer segments for more environmentally friendly practices [35]
Manufacturing focus	Lean manufacturing use Lean Automation [28]. Lean manufacturing focus on maximizing productivity by increasing output per unit of input, conserving resources, reducing waste, and minimizing costs [7]	Green manufacturing focuses on the environment and identifies waste as excessive use of resources or substances released to the air, water or land that could harm human health or the environment [24]
Product design focus	Product design sets the conditions for manufacturing: process and product development [18]; Eliminate waste, increase productivity and reduce cost [7]	Green design or Eco-design; Selection of low impact materials, reduction of materials usage, optimization of production techniques, improvement of distribution system, reduction of impact during usage, improvement of initial life time, and improvement of end of life system [18]
Inventory strategy	Minimizes inventory throughout the chain [4]	Reductions in inventory levels [4]
Supply focus	Inter-organizational involvement [22]. Close cooperation with all supply chain entities [18]	Inter-organizational collaboration involves activities to deploy green issues in supply chain entities [22]
Information sharing	Demand high levels of information sharing [22]	Information sharing is important for the successful implementation [22]
Employees	Employee involvement and empowerment [12]. The education and training of the employees to execute multiple tasks [2]	Involvement of employees [18]. The education of the employees concerning green issues [12]
Customer	Satisfying the customers by reducing costs and lead times [12]	Satisfying the customers by the implementation of green issues [12]
Processes	Focus on product [2]. Continuous improvement culture [24]	Focus on processes and products [18]. Process optimization through the implementation of green issues [31]

comparing the different characteristics between lean and green. Others important studies [3,4,12,18,22] were inspire for the development of a comparison between lean and green paradigms. Several lean and green supply chain characteristics are considered in Table 2.

4 Combining Industry 4.0 and Lean and Green Supply Chain

The future must become leaner in organization accurately in planning and technology [36]. The authors [21] mentioned that the paradigm Industry 4.0 solutions not only can be integrated in lean manufacturing but can be beyond that improve lean manufacturing. Also Sanders et al. [28] considered that industry 4.0 and lean manufacturing can be integrated to achieve a successful production management. However, they are not mutually exclusive [28].

Definitely, several researches mention the benefits of the integration of lean and green in different stages of the company or the supply chain [1,9,12,13,20]. Kainuma and Tawara [20] studied the lean and green supply chain incorporating there cycling or re-use during the life cycle of products and services. It represents the different phases of a product life cycle, consisting in [20]: (i) the acquisition of the raw material, (ii) the manufacturing, (iii) the distribution, (iv) the retailer, (v) the use, (vi) the collection, (vii) the transportation, (viii) the dismantling and (ix) the decomposition.

Stock and Seliger [30] presented the opportunities for the realization of a sustainable manufacturing in the Industry 4.0. For them the life cycle (in the end-to-end solution) consists in different phases [30]: (i) the raw material acquisition, (ii) the manufacturing, (iii) the transport (between all phases), (iv) the use and the service phase, and (v) the end-of-life phase (containing the reuse, remanufacturing, recycling, recovery and disposal). In addition, the environmental/green dimension of the sustainability is better considered because the allocation of resources as products, materials, energy and water can be realized in a more efficient way [30]. The adoption of smart energy systems facilitates the energy use [23].

In fact energy models would assist the analysis of green factory designs, especially for evaluating alternatives during early design stages [25]. The design of lean and green supply chain, special in the early design stages for the products and the processes is a very important issue for the elimination of waste. In a lean and green environment [7] mention that "eliminating the use of toxics through product or process re-design could mean reduced worker health and safety risks, reduced risks to consumers and lower risk of product safety recalls and reducing process wastes in manufacturing often find more opportunities to reduce waste throughout the life cycle of the product, thereby having a possible domino effect on the entire supply chain". Industry 4.0 is in line with these ideas. According to [27], Industry 4.0 processes will change the entire supply chains, from suppliers to logistics and to the life cycle management of a product. It helps to streamline the process, with more transparency and flexibility.

Lean and green supply chain requires manufacturing technologies to make processes and products more environmentally responsible [22]. In addition they ask for a flexible information system [17]. The technology is a driver of the Industry 4.0 [15,34]. With smart technologies which include the use of electronics and information technologies [27] will help the implementation of a more efficient lean and green supply chain.

Also the collaboration with suppliers which is a lean and green characteristic is considered by Industry 4.0. Through a better communication mechanisms, with a high compatibility issues of hardware and software which should required standardized interfaces, and synchronisation of data, allow that lean and green suppliers get better synchronisation with manufacturers [28,32].

The author [14] concludes that lean and green "is an effective tool to improve processes and reduce costs, by not only reducing non-value-added activities but also physical waste created by systems". Industry 4.0 is in line with this statement due this paradigm make all but in a better way, more sustainable, faster and efficient. According to [21] lean allows the organizations to be more standardized, transparent and having only the essential work which result in an organization less complex and support the installation of industry 4.0 processes and solutions. The green also support the implementation of the Industry 4.0 due it allows to reduce the negative environmental impacts.

The customer type is a concern in the lean and green supply chain. Of course that lean and green aim is to satisfy the customer needs, but this satisfaction is relative to: in the lean paradigm is based on cost and lead time reduction [9] and in the green paradigm is based on helping customers to being more environmentally friendly [13]. The Industry 4.0 will go to improve in this subject. It allows a better understanding of the customer needs and allows the immediate sharing of the demand data throughout complex supply chains [15]. According to [27] with the full automation and digitalization systems, it allows an individual customer-oriented adaptation of products that will increase the value added for organizations and customers. Customers instead of choose from a fixed product spectrum set by the manufacturer, they will be able to individually combine single functions and components and define their own product [15].

Another characteristic of lean and green supply chain is the employee involvement and empowerment [12]. According to [8] employee commitment and motivation, and employee empowerment and participation are elements of lean and green organization. Also [24] mention that connections between lean and green practices are shown through: (i) employee involvement, (ii) learning by doing, (iii) continuous improvement, and (iv) problem-solving tools. [36] mention that lean means reducing complexity, avoiding waste and strictly supporting the employees in their daily work. Also the reduction of environment impacts improves the health and safety of employees [31]. These aspects are in line of what is an employee in the four industrial revolution. Indeed, employees may find greater autonomy and more interesting or less arduous work [6]. Industry 4.0 needs employees not only with creativity and decision-making skills, encountered as a lean and green supply chain characteristic, as well as technical and ICT expertise [6].

There are in literature some studies that try to make the bridge between lean paradigm and green paradigm with Industry 4.0. [28] used 10 lean concepts in their research in way to validated for attainability through Industry 4.0 paradigm. Kolberg and Zühlke [21] described the lean automation and Industry 4.0 and give an overview of the links between them. [30] present an overview of sustainable manufacturing with the future requirements of Industry 4.0. Figure 1 illustrates an attempt to link lean and green supply chain characteristics to Industry 4.0 concepts.

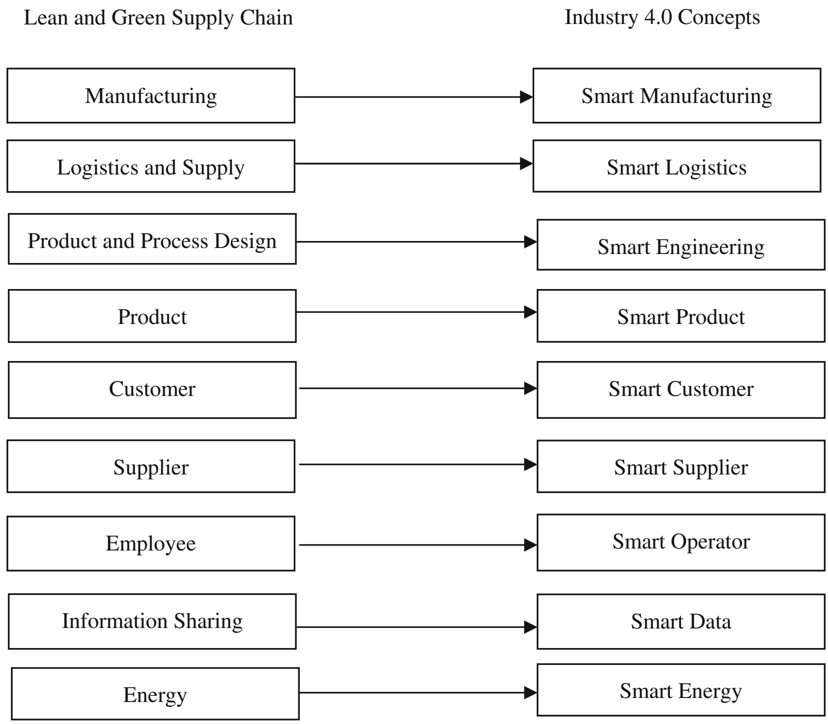


Fig. 1. Linking the lean and green supply chain characteristics to the Industry 4.0 concepts

5 Conclusion

Today, the term Industry 4.0 describes a vision of future of the supply chains. There is a strong conviction that the definition of lean and green supply chain will not disappear, it will be evolve and adapt to the new trends that the new industrial era will require. Lean and green supply chain is focussed on organization and in the flow of information, material and money between partners. That is, more directed to physical processes and less for virtual and technology. Even so there are in literature some examples that try to make the bridge between lean paradigm or green paradigm with Industry 4.0. This paper bridges the gap between the well-known lean and green supply chain management and the new era of industrial revolution.

A conceptual model was developed linking the lean and green supply chain characteristics to the Industry 4.0 concepts. Several characteristics were presented on model, namely: (i) manufacturing, (ii) logistics and supply, (iii) product and process design, (iv) product, (v) customer, (vi) supplier, (vii) employee, (viii) information sharing and (ix) energy. Those who understand the relationships between these two topics will have a greater chance of influencing their supply chains into a source of competitive advantage and help in a better way on the deployment of the Industry 4.0 paradigm.

Future research is needed. Understand which lean and green characteristics are more important for the development of Industry 4.0 is required. It would be also beneficial to understand the priority between characteristics on the implementation of this new paradigm and in different entities in the supply chain. Industry 4.0 will be a step forward for the effectiveness and competitiveness of the lean and green supply chains.

Acknowledgements. Authors would like to acknowledge to UNIDEMI–Research and Development Unit for Mechanical and Industrial Engineering (UID/EMS/00667/2013).

References

1. Azevedo SG, Carvalho H et al (2012) Influence of green and lean upstream supply chain management practices on business sustainability. IEEE Trans Eng Manage 59(4):753–765
2. Bortolotti T, Boscari S, Danese P (2015) Successful lean implementation: organizational culture and soft lean practices. Int J Prod Econ 160(12):182–201
3. Carvalho H, Azevedo SG, Cruzmachado V (2010) Supply chain performance management: lean and green paradigms. Int J Bus Perform Supply Chain Model 2(3):304–333
4. Carvalho H, Duarte S, Machado VC (2011) Lean, agile, resilient and green: divergencies and synergies. Int J Lean Six Sigma 2(2):151–179
5. Carvalho H, Azevedo S, Cruz-Machado V (2014) Trade-offs among lean, agile, resilient and green paradigms in supply chain management: a case study approach. Lect Notes Electr Eng 242:953–968
6. Davis R (2015) Industry 4.0, digitalisation for productivity and growth. European Parliamentary Research Service (EPRS), Members' Research Service, European Union
7. Dhingra R, Kress R, Upreti G (2014) Does lean mean green? J Cleaner Prod 85:1–7
8. Duarte S, Cruz-Machado V (2013) Lean and green: a business model framework. Lect Notes Electr Eng 185:751–759
9. Duarte S, Cruz-Machado V (2013) Modelling lean and green: a review from business models. Int J Lean Six Sigma 4(3):228–250
10. Duarte S, Cruz-Machado V (2015) Investigating lean and green supply chain linkages through a balanced scorecard framework. Int J Manage Sci Eng Manage 10(1):20–29
11. Duarte S, Machado VC (2011) Manufacturing paradigms in supply chain management. Int J Manage Sci Eng Manage 6(5):328–342
12. Dües CM, Tan KH, Ming L (2013) Green as the new lean: how to use lean practices as a catalyst to greening your supply chain. J Cleaner Prod 40(2):93–100
13. Garzareyes J (2014) Lean and green-synergies, differences, limitations, and the need for six sigma. IFIP Adv Inf Commun Technol 439:71–81
14. Garzareyes JA (2015) Green lean and the need for six sigma. Int J Lean Six Sigma 6(3):226–248
15. Germany Trade & Invest (GTAI) (2014) Industry 4.0-Smart Manufacturing for the future. Germany Trade and Invest

16. Hines P (2009) Lean and Green, 3rd edn. Sapartner, New York
17. Jasti NVK, Kodali R (2015) Lean production: literature review and trends. Int J Prod Res 53(3):867–885
18. Johansson G, Winroth M (2009) Lean vs. green manufacturing: Similarities and differences. In: International euroma conference: implementation - realizing operations management knowledge
19. Kagermann H, Helbig J et al (2013) Recommendations for implementing the strategic initiative german industrie 4.0. final report of the industrie 4.0 working group. Technical report, Forschungsunion
20. Kainuma Y, Tawara N (2006) A multiple attribute utility theory approach to lean and green supply chain management. Int J Prod Econ 101(1):99–108
21. Kolberg D, Zühlke D (2015) Lean automation enabled by industry 4.0 technologies. IFAC Papers Online 48(3):1870–1875
22. Mollenkopf D, Stolze H et al (2010) Green, lean, and global supply chains. Int J Phys Distrib Log Manage 40(1/2):14–41
23. Noppers EH, Keizer K et al (2016) The importance of instrumental, symbolic, and environmental attributes for the adoption of smart energy systems. Energy Policy 98:12–18
24. Pampanelli AB, Found P, Bernardes AM (2014) A lean & green model for a production cell. J Cleaner Prod 85:19–30
25. Prabhu VV, Jeon HW, Taisch M (2012) Modeling green factory physics - an analytical approach. In: IEEE international conference on automation science and engineering, pp 46–51
26. Qin J, Liu Y, Grosvenor R (2016) A categorical framework of manufacturing for industry 4.0 and beyond. Procedia Cirp 52:173–178
27. Roblek V, Meško M, Krapež A (2016) A complex view of industry 4.0. Sage Open 6
28. Sanders A, Elangeswaran C, Wulfsberg J (2016) Industry 4.0 implies lean manufacturing: research activities in industry 4.0 function as enablers for lean manufacturing. J Ind Eng Manage 9(3):811
29. Shrouf F, Ordieres J, Miragliotta G (2014) Smart factories in industry 4.0: A review of the concept and of energy management approached in production based on the internet of things paradigm. In: IEEE international conference on industrial engineering and engineering management, pp 697–701
30. Stock T, Seliger G (2016) Opportunities of sustainable manufacturing in industry 4.0. Procedia Cirp 40:536–541
31. Verrier B, Rose B, Caillaud E (2015) Lean and green strategy: the lean and green house and maturity deployment model. J Cleaner Prod 116:150–156
32. Veza I, Mladineo M, Gjeldum N (2016) Selection of the basic lean tools for development of croatian model of innovative smart enterprise. Tehnicki Vjesnik 23(5):1317–1324
33. Vonderembse MA, Uppal M et al (2006) Designing supply chains: towards theory development. Int J Prod Econ 100(100):223–238
34. Wang S, Wan J et al (2016) (2016), Implementing smart factory of industrie 4.0: an outlook. Int J Distrib Sensor Netw 4:1–10
35. Zhu Q, Sarkis J, Lai KH (2008) Confirmation of a measurement model for green supply chain management practices implementation. Int J Prod Econ 111(2): 261–273
36. Zuehlke D (2010) Smartfactory-towards a factory-of-things. Annu Rev Control 34(1):129–138

A Model of Maker Education in Chinese Universities: The Perspective of Innovation Ecosystem

Qinglong Zhan[1(✉)] and Mengjia Yang[2]

[1] School of Information, Tianjin University of Technology and Education, Tianjin 300222, People's Republic of China
qlzhan@126.com
[2] Tianjin Nankai High School, Sino-Singapore Tianjin Eco-city, Tianjin 300222, People's Republic of China

Abstract. Maker education is an important mission of Chinese universities for the entrepreneurship and innovation. However, current maker education failed to meet the makers' demands for real-time awareness of space, adaptive automatic resources, diverse innovation activities and tools with intelligent support in the makerspace. Therefore, the innovation ecosystem has emerged to be the best choice to solve those problems. This paper first reviews literatures on maker education, and then analyses fundamental principles and elements of the innovation ecosystem. After that, essentials and properties are discussed for innovation ecosystem-oriented maker (IEOM) education. At last, the model of IEOM education is developed, including innovation subject, innovation activity, and smart innovation environment.

Keywords: Maker education · Innovation ecosystem · Chinese universities

1 Introduction

Maker education is an important mission of Chinese universities for the entrepreneurship and innovation. However, current maker education in Chinese universities has the limitations of closed innovation, efficiency-driven, focal product, specialized organization, and so on. They use a hierarchical management team as the subject of innovation, and use virtual laboratories, heavyweight wearable devices and other technologies to build an intelligent and well-defined physical space. This failed to meet the makers' demands for real-time awareness of space, adaptive automatic resources, diverse innovation activities and tools with intelligent support in the makerspace. It also does not comply with their hopes of cooperative innovation, wisdom communications and results-sharing. Therefore, the innovation ecosystem has emerged to be the best choice to resolve

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_104

those problems. At the same time, it needs to construct a model of innovation ecosystem-oriented maker (IEOM) education as an important accelerator of the sustainable development of university makerspace. In IEOM education, any maker can be innovative as long as there are a true interest and consistency in establishing a favorable environment to support it [4].

In this paper, we first review literatures on maker education, and then analyses fundamental principles and elements of the innovation ecosystem. After that, essentials and properties are discussed for innovation ecosystem-oriented maker (IEOM) education. At last, the model of IEOM education is developed, including innovation subject, innovation activity, and smart innovation environment.

2 Literature Review

Maker education of universities arises from the maker movement with the learning by doing culture. The maker movement has provided a perspective on learning that differs from the traditional learning practices taking place in schools and universities [14]. Barrett, Pizzico, Levy and Nagel [1] discussed the benefits of university makerspaces that are primarily focused around building physical models and the inherent of informal learning environments and community. Physical models can increase the effectiveness and quality of the final design for the development of undergraduates by linking the material covered in the classroom to the real world. Informal learning environments are more open environment allows them to be used more freely and interwoven into the class structure for multiple classes without typical classroom scheduling constraints.

The best practices of university makerspace are investigated by several researchers. Wilczynski [23] explored academic maker spaces in universities, such as Arizona State University, Georgia Institute of Technology, Massachusetts Institute of Technology, Northwestern University, Rice University, Stanford University, and Yale University. He recommended the best practices as follows: The mission of the academic makerspace must be clearly defined; Ensure that the facility is properly staffed; Open environments promote collaboration; Aligning access times with the student work schedules; Providing user training is essential; Attention must be devoted to establish a maker community on campus. Katona, Tello, et al. [10] put forwards four best practices: flexible models for interdisciplinary faculty hiring and engagement, development of student entrepreneurs, integrating cross-campus curricula, and the development of cross-campus collaborations. Myers [12] encouraged interdisciplinary research and idea, which elements were student led engagement, access to the latest technology, and key partnerships. Bieraugel [2] found that the on-campus makerspace located outside the university library encouraged the most innovative behaviors and exploration of new ideas, and within the library, collaboration rooms were the best spaces for encouraging creativity.

As to Chinese researchers, their focuses are: relationships between maker education and innovation education [8,25]; Essentials and functions of maker education [27–29], double helix model [24] and design-based learning [9,30]; augmented virtual technology and Internet plus in maker education [19,20,27], development paths of maker education [18,21] and system framework [26].

Obviously, innovation ecosystem that can be aligned with university maker education currently seems to be not involved in the literatures. In fact, it is a good idea to solve problems that development of maker education faces in Chinese universities.

3 Theoretical Roots

The innovation ecosystem is a newer and mainstream concept discussed widely by scholars and practitioners. However, innovation ecosystem is often described in different ways, and thus understanding it remains a challenge. Luomaaho and Halonen [11] defined innovation ecosystem as a permanent or temporary system of interaction and exchange among an ecology of various actors that enable the cross-pollination of ideas and facilitates innovation. Nordfors [13] referred that innovation ecosystems embody technology and information flow between those needed to turn ideas into processes, products or services. Bulc [4] pointed out that an innovation ecosystem is a system made for innovation creation in an open, natural manner, which enables a holistic understanding of needs, solutions, and consequences related to innovation processes and innovation itself, and is an interaction between people, enterprises, and institution. Ritala [15] had a view that innovation ecosystem refers to clusters (physical or virtual) of innovation activities around specific themes (e.g., biotechnology, electronics and software). The complex innovation ecosystem where networks of innovations and communities of people and organizations interact to produce and use the innovations [22]. Taken all together, Innovation ecosystem is an interactive work environment which strengthens actors of different disciplines to co-create new ideas, and then to construct prototypes and commercialize them in a way of ongoing innovation.

An innovation ecosystem has several important elements. The basic components of the ecosystem must be a high quality (universities, funding possibilities, specialized services, talent pool and regional dynamics) [7]. Bulc [4] thought that key elements can be put into 4 major groups: participants, tools, content, and principles. Haines [6] identified the following key ingredients of an innovation ecosystem: culture, champion(s), network, stakeholder engagement, process, physical space and events. Obviously, the innovation ecosystem includes funds, material resources (equipment or facilities), and the innovative actors such as students, teachers, researchers, industry specialists, venture capitalists. They participate in the ecosystem to enable technology development and all kinds of innovations.

Creating an innovation ecosystem needs to think about some basic principles: getting rid of the box, core value, competences, dynamic structure, and constant change [4]. To make innovation, a suitable innovation ecosystem must meet different conditions such as natural, structural, organizational and cultural factors. Those factors can be grouped based on the following dimensions: resources, governance, strategy and leadership, organizational culture, human resources management, people, partners, technology and clustering [5]. A flow of technology and information is the key factor among the actors: the people, enterprises, and institutions [16].

4 Essentials and Properties of IEOM Education

4.1 Essentials of IEOM Education

IEOM education creates a new paradigm for a large-scale cooperative design, creation, invention, and innovation. It develops a novel way of intellectual association, transparent operation, and collaborative participation in the innovation. In this novel approach, creative ideas, services, and products can be implemented, and the greatest value can be achieved.

IEOM education takes innovative actors as a community and utilizes new technologies to create a real world of participatory innovation. These technologies are emotional computing, open source hardware, holographic environment, 3D printing, social media and adaptive technologies, etc. IEOM education is not only the emergence of new technologies but also the emergence of transformative innovation activities participated by makers. Firstly, due to the implementation of the open source hardware and open resources, IEOM education provides multiple paths involved in innovation activities for makers full of creative thinking and creativity. Secondly, makers can not only connect with people, machines and objects to the physical space, but also seamlessly connect all the information sources outside the space. Makers' creative partners are also just from members of a team to global actors in the project.

4.2 Properties of IEOM Education

(1) Wisdom circle
The wisdom circle is proposed to be a new concept in the development of the biosphere, emphasizing the powerful forces of human wisdom in modern evolution. Human beings give full play to collective wisdom to explore new methods, to create innovative technologies, and to develop new tools. They use created products to change the makerspace, further enhance the intelligence of the group, expand the scope of the role. Ultimately, the makerspace evolves into the human wisdom circle containing capacity for development and creativity. The wisdom circle relies on the rapid flow and superior connectivity of data, information and wisdom. The flow and connectivity produce great transformation power, and establish a fresh creative field of relating to wisdom individuals, enhancing group wisdom and activating participatory innovation experiences.

(2) Micro-innovation
While micro-innovation is characterized by a small team, it shares the common trait with macro innovation: creative thinking based on recurring actions, re-engineering an opportunity to make every thought new, and building creative solutions that bring true progress. Micro-innovation is a local action and a

global innovation. It is a decentralized, network-based innovation. Combinations of indigenous micro-innovation will have a direct impact on the overall innovation. Micro-innovation is not holistically planned and guided. In this sense, the innovation is a daily habit. Micro-innovation encourages new technologies and processes to be extended in IEOM education, and offers sustainable innovative results for any size of the makerspace.

(3) Cross-border
As a creative scene, the makerspace in IEOM education is cross-organizational, inter-spatial and interdisciplinary. It breaks the closed boundaries in order to truly achieve the integration of innovation. Firstly, the makerspace is not only beyond the visible border of physical space, but also is across the invisible boundary of the digital space. Secondly, the wisdom flow and the information flow continually mobilize among the enterprises, research groups, project groups, innovative communities and other organizations in IEOM education. Thirdly, the acquisition of innovation is not confined to a specific discipline. Therefore, the IEOM education emphasizes cross-border, connection and integration among the different disciplines and spaces.

(4) High value-added
There are ongoing associations, high circulation and self-enhancement in IEOM education. It is the circulation and re-creation of data and information, and environmental innovation that produces the most valuable knowledge. For those data and information of weak circulation and low re-creation, it is not able to create leapfrog development and high valued innovation. Therefore, data is needed to be reprocessed deeply, and the knowledge flow is needed to speed up. Further to this, individual wisdom and the power of group innovation are converged to achieve a high value-added and an upward spiral of innovation capabilities showing the development of creativity. For example, scientific knowledge and big data generated by activities, and deeper structure for mining methods, are combined with language understanding and image recognition to provide value-added services for makers.

(5) Strong lean
IEOM education of strong lean emphasizes scientific analysis of big data in the creative process, and optimizes decision-making. Strong lean can improve the performance and reduces the bias, with the aim of discovering potential problems, assessing the progress of creation, and predicting future performances. Among them, lean thinking is scientific and accurate, thereby being efficient and optimal. In strong lean, analysis technologies of data and information visualization can be used to present the results of complexity analysis such as charts or stereoscopic images. In doing so, relationships and patterns of things are identified quickly and easily.

(6) Multi-interactions
Multi-interactions include internal interactions between human-to-thing (H2T), human-to-machine (H2M), thing-to-machine (T2M), human-to-human (H2H),

thing-to-thing (T2T). They provide more maker-oriented services for IEOM education and meet the maker's strong demands of highly interactive experiences. The friendly multi-interactions are inseparable for designing human-thing communication devices, machine operation objects, and human-computer interaction. It provides natural, convenient and efficient interactions among makers, machine and things.

5 Design Principles of IEOM Education

IEOM education takes Web 4.0 as the basis and creation as being the engine to establish the innovative ecosystem of the overall situation. Its mission is to create an innovative ecosystem that integrates the various innovation activities into IEOM education. The core is to create synergies among the innovative collaborators, such as industries, educational institutions and research institutions. The paradigm is an innovative network which is composed of innovative actors.

IEOM education focuses on new technologies and innovation strategies. It also concentrates on innovators, new technology adoption, cross-cutting issues, learning organizations, forms of activity, and innovative business environments. IEOM education balances process innovation and creativity by establishing innovative space, and builds a network of innovative culture and a collaborative innovation model. Thus by strengthening the creative culture and interdisciplinary innovative solutions, IEOM education brings the integration of innovation strategy and digital technology to an unprecedented level.

IEOM education produces a combination of innovative technologies and provides better user experiences of contextualization and value-added, supported by digital technologies such as social media, mobile computing, big data, cloud computing, sensor networking, Internet and machine intelligence. For example, social media networks enhance coexistence; big data provides a unique analysis, targeting the recommendations and preferences in the situations; mobile technology creates location context services and awareness; cloud computing enlarges access to resources and services; sensors create real-time feedback and response; machine intelligence improves the quality of insight and decision-making.

The digital space is a new method to drive the innovative wisdom of individuals, organizations, and communities to establish a network of cross-industry and actors' innovation. It expands the ecosystem leading to innovation and creation. Digital space in IEOM education connects various things and the actors to enhance the relevance and importance of digital technology. Through newly connecting and integrating digital technologies aforementioned, the digital space can enhance collaboration, and access the results of high value-added.

With the integration of digital technologies, IEOM education takes Internet into the real-world, to form the human networking, things networking, and service networking. This integration can configure technology, media, market and actions of the actor altogether, and cross-links smart applications of real-world products, services, and location by Web 4.0.

6 Building a Model of IEOM Education

IEOM education consists of three elements of innovation subject, innovation activity, and innovation environment. Innovative subjects include experiencers, innovators, researchers and experts, namely, the inner creative knowledge space. Innovation activities are developed by the innovation subjects including interest, research, practice and smart creation, namely, external creative knowledge space. The innovation environment is a creative scenario of a whole new perspective, reflecting innovative subjects' rich creation in participating in innovation activities. The model of IEOM education is shown in Fig. 1.

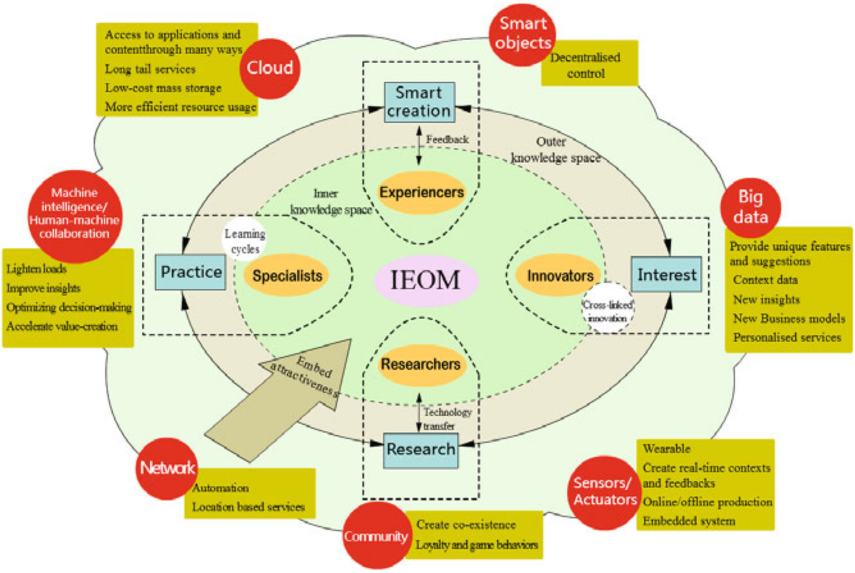


Fig. 1. Development model of IEOM education

6.1 Innovation Subjects

(1) Experiencers
The experiencers are the subjects of intellectual creation, are the explorers of the unknown world, and are the perceivers of facing the future. On the basis of light capital and low investment, numerous experimenters learn, experience and absorb in IEOM education, and then redesign and transform knowledge and activities. In IEOM education, experiencers affect the development of the real industry remodeling the physical world, to generate digital, open and popular desktop manufacturing industry. Experiencers take the dexterous participation and agile action to input new perceptions and ideas for the intellectual creation activities. At the same time, intellectual activities also provide feedback on innovative activities for the experiencers.

(2) Innovators
In IEOM education, innovators with the similar interests gather together to try unique combinations of multi-domain, multi-entity, multi-method and cross-linking innovations. They blur boundaries, continuously redefine, and constantly remodel works. Innovators connect IEOM education to nature and life in order to design a unique perspective of the new works, such as technology + nature + art, digitization + virtualization + audiovisual performances, and other new explorations.

(3) Researchers
Facing the unknown future, researchers use systematic thinking and scientific methods to deepen the essence of things and the main contradictions. They capture new knowledge, generate new ideas, utilize new technologies, and explore new approaches. Rather than the reality of unchanging model, researchers study the possibility of the future to help the development with the IEOM education. It is important for researchers to migrate and remix technologies, which are an evolution path of new technologies. Similarly, the study of innovative activities can also stimulate the creative thinking and technological innovation of the researchers.

(4) Experts
Experts emphasize the professionalism of their subject knowledge and the sensitivity to complex things. They are not to solve problems according to the steps, but to focus on asking questions, mining data, clarifying the relationship and finding a solution. In IEOM education, the experts transcend over-fitting (self-perceived vertexes) through cross-border connectivity and cyclic learning. At the same time, on the basis of strong information filtering, experts show innovative wisdom of research on product-updating, platform-upgrading, and entity-high-value-reusing.

6.2 Innovation Activities

(1) Interest
Interest is the engine of makerspace development. Innovators gather in the makerspace, generate creative ideas and innovative products in a more personalized and independent way, based on aspirations, visions and interests. In IEOM education, interest-based innovation activities, are not a clear bounded knowledge or experience passed to the creators, but a creative enthusiasm for innovation, social responsibility and other elements to affect makers. Innovation activities emphasize cross-border, connectivity, collaboration and sharing, to enhance the cohesion of makers' innovations. They highlight the motive of exploration and innovation, to encourage self-learning, decision-making and creativity. Innovative activities show the powerful forces of innovative self-development and co-creation through the de-centralization, peer organization. They offer high-yield innovations with substantial value-added.

(2) Research
Research is the foundation of the development of IEOM education. Through the unknown, undetermined, complex, chaotic and other advanced technology exploration and the transfer of key technologies, research achieves changes and technological breakthroughs. As for the entities, products, and services, a researcher can flexibly design, massively customizes and agilely develops. The aspects of design, construction, and testing may be updated or changed iteratively.

(3) Practice
Makerspaces are central resource spaces offering various making practices in different contexts. Practice is an acceleration power of innovation. Practice is the only way to produce knowledge creation and innovation. Creative activity in practice refers to the fact that the experts absorb, interact, reflect, criticize and recreate knowledge and products through the context-based collection, correlation and integration. It stresses the unity of theory and practice, and shows the great wisdom of innovation and creation. In IEOM education, cross-domain experts carry out innovative ideas and actions to the practice of product development. This practice appears in a circle of knowledge production and application transformation.

(4) Smart creation
In industry, smart creation means to quickly and adaptively fabricate products using automatic, networked and intelligent information technologies such as Internet of things. In IEOM education, smart creation can be all a matter of structural change, production optimization, efficiency enhancements and value-added. It is decided on core skills, intellectual capital and other top factors of the practitioners. Smart creation is the dexterous behavior, which the individual and the collective in the creative network culture focus on the tiny creativity, the agile product development, the custom individuality service, and the creative behavior reshaping. Smart creation demonstrates the interaction and collaboration between makers by enhancing their dexterity behavior and the creativity practice. With IEOM education, it increases the possibility of the creative exploration, scientific attempt and innovative practice for makers.

6.3 Innovation Environments

(1) Cloud services of IEOM education
Hyper-scale data gathered in a creative environment is inseparable from the bearing and calculation capacity provided by the cloud. Cloud supports the integration and sharing of large-scale resources and has a low requirement on equipment, multi-channel access to resources, low-cost storage of large-scale data, and efficient use of data resource to deliver long-tail stable service. Cloud has features of on-demand access, ubiquitous access, resource pooling, elasticity, measurable use, recoverability, high security and high scalability. These features of cloud computing are essential to the reality of demand for IEOM education.

(2) Smart objects of IEOM education
As the cognized smart objects, they empower objects or things corresponding cognitive abilities and reflect decentralized control features of the non-linearness, openness and flatness. As an entity embedded in the autonomously distributed intelligent system, smart objects can intelligently act on the human-machine-thing objects based on the sensor's perception in IEOM education. Multi-object cooperation is able to globally optimize, to percept real laws and to solve the major complexity of the problem by a distributed cooperative work.

(3) Big data of IEOM education
Big data in IEOM education aims at all the context data generated by the innovation subject in the innovation activity. Through batch data processing and unique perspective analysis of all the relational and non-relational data, it provides accurate forecast and personalized real-time suggestions for the innovation subject. It can be a keen insight into the quality of data, data-mining in the hidden information, and the true reflection of the state of activity. At the same time, weak related data floods with sparse value density can still be appropriately associated and deeply mined, in order to create a new situation and a new development stage of IEOM education.

(4) Sensors of IEOM education
Sensors provide large amounts of real metadata for big data by detecting objects, observing events, tracking targets, describing or quantifying activities and capturing information. In IEOM education, the main function of sensors is perception and cognition. The perception is the fact that the different sensing elements perceive the context to obtain the parameters of the quantized and non-quantized. Cognition is tantamount to acquire the state information or knowledge, such as subjective emotion, psychological state, and another individual implicit information. Cognition can abstract the relationship of information recorded in the communication, properties and location of devices and tools, and others. Sensors own the functions of self-capturing, self-tuning, self-identification, and logical judgments. They provide on-line or off-line production and establish real-time information feedback, moving towards the depth of miniaturization and high-level intelligence.

(5) Communities of IEOM education
Members of a community are technologically sophisticated do-it-yourself (DIY) enthusiasts interested in activities related to engineering, electronics, robotics, and desktop fabrication [3]. In the inter-connected times, the community reconstructs the circle of interaction, re-establishes the role of identity, reconstructs the cooperative relationship, and creates the group relationship of real and virtual coexistence. The community is connected. The connection is namely development. In IEOM education, the innovative subjects connected with interest will establish the creative community and is involved in the innovation activities. This community establishes a co-existence and co-construction, and shares in loyalty or game behaviors.

(6) Network of IEOM education
IEOM education networks provide links be-tween specific makerspace workshops and projects that permit forms of identities to flow and mobilize [17]. Innovation networks in IEOM education give innovation subjects insights into temporal and spatial contexts, in which to capture preferences, goals, paths, activities, and locations. They can automatically connect appropriate information, resources and nodes with humanized supports. Location-based networked services are not just based on the physical space of the subject, but also the mind and emotional states displayed by the subject of activities.

(7) Machine intelligence and human-machine collaboration in IEOM education
IEOM education can be inextricably linked to cyber-physical systems (CPS), which is the in-depth integration of computation, communication, and control technologies. CPS couples network and computing to the physical environment, and to achieve full autonomy, efficient collaboration, mobile control, dynamic adaptation, and human services. On the one hand, human-machine collaboration transforms the interaction between human and machine and object, and generates new methods. On the other hand, it associates individual objects or subsystems into a unified entity, setting up a fusion of the natural world, the virtual world, and the thought world. CPS emphasizes the communication between human and machine, and scientifically controls the movement of objects through precise calculation of data parameter. These communications and controls are woven into innovation activities and continuously facilitate innovative subjects sensing objects and knowing laws.

7 Conclusion

IEOM education not only drives the development of the world but also the progress of humanity. It transforms the external world, while at the same time shaping the inner world of mankind which conforms to the characteristics of the new era. IEOM education takes individual intelligence to be gathered into a powerful force, and to create a new situation of humanization, open-sources, and co-creation. It breaks the limitations of traditional innovation and closed research, to change people's ideas and thoughts in the traditional innovation environment, and promotes the development of the makerspace.

Acknowledgements. This work is supported by a grant from the Research Program of Science and Technology Development Strategy of Tianjin "Research on Mechanism and Model of Innovation Ecosystem-oriented Makerspace in Tianjin universities" (No. 16ZLZXZF00410).

References

1. Barrett T, Pizzico M et al (2015) A review of university maker spaces. Georgia Institute of Technology

2. Bieraugel M, Neill S (2017) Ascending bloom's pyramid: fostering student creativity and innovation in academic library spaces. Coll Res Libr 78:3–52
3. Brown A (2015) 3D printing in instructional settings: identifying a curricular hierarchy of activities. TechTrends 59(5):16–24
4. Bulc V (2011) Innovation ecosystem and tourism. Academica Turistica-Tourism Innov J 4(1):27–34
5. Durst S, Poutanen P (2013) Success factors of innovation ecosystems. http://www.academia.edu
6. Haines T (2016) Developing a startup and innovation ecosystem in regional Australia. Technol Innov Manage Rev 6(6):24–32
7. Hautamäki A (2006) Innovation ecosystem in city policy: the case of Helsinki. http://www.edu.helsinki.fi
8. He K (2016) On maker education and innovation education. Educ Res 4:12–24 (in Chinese)
9. Huang L, Bao X, Wang Y (2016) Learning by design: a new study on maker education practice model in schools. China Educ Technol 11:18–22 (in Chinese)
10. Katona T, Tello S et al (2015) Pathways partners: entrepreneurial change across campus. In: VentureWell. Proceedings of open, the annual conference, national collegiate inventors & innovators alliance, p 1
11. Luoma-aho V, Halonen S (2010) Intangibles and innovation: the role of communication in the innovation ecosystem. Innov Journalism 7(2):3–20
12. Myers J (2015) Creating collaborative spaces at the university of Arizona: ways to encourage interdisciplinary research and ideas
13. Nordfors D (2009) Innovation journalism, attention work and the innovation economy. A review of the innovation journalism initiative 2003–2009. Innov Journalism 6(1):5–46
14. Pont E (2015) Analysis and comparison of representative locations in the general makerspace panorama. http://upcommons.upc.edu/handle/2117/87120
15. Ritala P, Agouridas V et al (2013) Value creation and capture mechanisms in innovation ecosystems: a comparative case study. Jpn J Appl Phys 63(3–4):244–267
16. Schwartz D, Barel R (2015) The role of a local industry association as a catalyst for building an innovation ecosystem: an experiment in the state of ceara in Brazil. Innovation 17(3):383–399
17. Smith A (2016) Grassroots digital fabrication and makerspace. https://ideas.repec.org
18. Tian Y (2016) Maker education: its origin, connotation and possible paths. Int Comp Educ 38(1):22–28 (in Chinese)
19. Wan L, Kang C (2016) Internet plus maker education: construction of the new ecology for innovation and entrepreneurship education in universities. Res Educ Dev 7:59–65 (in Chinese)
20. Wang D, Song S, Chen Z (2016) Application of virtual reality and augmented reality in maker-driven education in universities. China Educ Technol 10:112–115 (in Chinese)
21. Wang M, Zhang L, Zheng Z (2016) The implementation approach of maker education in our country. Mod Educ Technol 26(9):11–17 (in Chinese)
22. Wang P (2009) An integrative framework for understanding the innovation ecosystem. http://doku.iab.de
23. Wilczynski V (2015) Academic maker spaces and engineering design. Am Soc Eng Educ 26:1

24. Yang G (2016) The construction of the double helix model of maker education. Mod Distance Educ Res 139:62–68 (in Chinese)
25. Yang G (2016) Maker education: the new path to the development of creative education in China. China Educ Technol 3:8–13 (in Chinese)
26. Yang L, Zhang L, Wang G (2016) Study on the system framework of maker education. Mod Distance Educ 3:28–33 (in Chinese)
27. Zhan Q, Yang M (2015) Research on maker education 2.0 and smart learning activities from the perspective of "internet plus". J Distance Educ 6:24–31 (in Chinese)
28. Zhang M, Liu X, Zhang C (2016) Connotation, function and reflection of maker education. Mod Educ Technol 26(2):14–19 (in Chinese)
29. Zhong B (2016) Discussion on background, essence, form, and support system of maker education. Mod Educ Technol 26(6):13–19 (in Chinese)
30. Zhu L, Hu X (2016) Research on maker education oriented design based learning: model and case. China Educ Technol 358:23–29 (in Chinese)

Pricing and Greening Policy for Manufacturer with Low Carbon Strategic Customers

Wen Jiang(✉)

College of Architecture and Urban-Rural Planning,
Sichuan Agricultural University, Chengdu 611830, People's Republic of China
xuezhongsha_wen@163.com

Abstract. Global warming is mainly caused by excessive emissions of greenhouse gases especially the carbon dioxide, which makes the carbon reduction become hot issues to businesses and consumers. This article considers a manufacturer produce one product and assumes the customers are of low carbon and consisting of myopic customers and strategic customers. The customer decision is affected by the price and unit carbon emissions of the product. The paper examines the optimal pricing and greening policies for the manufacturer with and without green technology investment. It shows that the optimal prices in two scenarios are decreasing in the ratio and patience of strategic customers. The presence of strategic customers is detrimental to the manufacturer and the negative impact increases as the ratio and patience of strategic customers increases. It also shows that with the addition of carbon sensitive degree, the manufacturer will improve the regular sale price and invest more in green technology when the customer number exceed a certain threshold. Finally, the article analyzes the effect of green technology investment on the optimal price and find that the optimal price with green technology investment is lower than that without green technology investment.

Keywords: Strategic customer · Carbon sensitive demand · Pricing decision · Green technology investment

1 Introduction

Global climate change is a serious threat to human survival and development. Global warming is mainly caused by excessive emissions of greenhouse gases especially the carbon dioxide [9], which makes the carbon reduction become hot issues to businesses and consumers [5,10]. In response to the Government's carbon reduction requirements and low carbon preference of customers, manufactures are considering whether invest in green technologies to reduce the carbon emissions of their products. With the implementation of the carbon footprint, it is becoming easier for customers to obtain the information on carbon emissions and to take this information into account in purchasing decisions. Therefore, it is very meaningful to study the optimal decisions of the manufacturer with carbon emissions sensitivity demand. Another phenomenon worthy

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_105

of study is the presence of strategic customer behavior [7]. With the development of e-commerce and the application of dynamic pricing strategies, customers are increasingly able to anticipate future price cuts and chose to delay the purchasing. Such "smart" customers are called strategic customer and customers who will not anticipate discounting are called myopic customer. The strategic and myopic customers are co-exist in real-life. Firms who ignore the presence of strategic customer will result in 20% losses of their total profit [1].

Two research streams are closely related to this article and will be reviewed to highlight our contributions. In the literature of models with carbon emissions demand or green technology investment, Bansal and Gangopadhyay [2] found that consumers were willing to pay a higher price for low-carbon products. Klingelhö [11] examined the effect of carbon emissions trading on green technology investment. The paper found that taking carbon emissions trading into consideration had impact on the investment, but this impact was not always positive incentives. Zhao et al. [17] investigated how did the manufacture in perfectly competitive market chose a technology when green technology was available. Yalabik and Fairchild [16] examined the manufacturer's optimal green technology investment of environment-friendly products with consumer choice was taken into consideration. Xia et al. [15] analyzed the interrelations among the carbon emission reduction, green technology investment choices and enterprise performance. These study provided a reference for the study of this paper, but they all assumed the customers were myopic and ignored the fact that myopic and strategic customers co-exists in real-life. In addition, Su [14] examined the pricing model of a monopolist who sold finite inventory to heterogeneous strategic customers in finite period. Jerath et al. [6] investigated the impact of two sales approach (through an opaque intermediary versus last-minute sales directly to consumers) on the service provider's profit when strategic customer behavior was considering. Lai et al. [12] investigated the effect of a posterior price matching policy on the customers' purchasing behavior, and seller's pricing decisions. They showed that strategic consumers' waiting incentive could be eliminated by this policy. Du et al. [4] examined the single-period joint inventory and pricing strategies with strategic customers and risk preference. Prasad [13] considering mix of myopic and strategic consumers and studied the choices of mixed bundling pricing and reserved product pricing for a firm who sold two products. Jiang and Chen [8] studied the optimal decision for a low carbon supply chain consists of one supplier constrained by cap-and-trade policy and one retailer facing homogeneous strategic customers. The literature above does not consider carbon emissions, or assumes that customers are homogeneous.

In order to fill the gap present above, this paper examine the manufacturer's pricing and green technology investment policy when the customer are carbon emissions sensitive and myopic and strategic hybrid. This paper aims to address following three questions:

- What is the optimal pricing policy for the manufacture without green technology investment? How does the presence of strategic customers affect the optimal pricing?
- What is the optimal pricing and greening policy for the manufacture with green technology investment? How does the presence of strategic customers affect the optimal pricing and greening policies?
- What effect does the green technology investment on the optimal policies?

2 Model Descriptions and Assumption

The paper considers one manufacturer who produce a product and direct sell it to the end users. The whole sales period is divided into two stage. At the first stage, the manufacturer sells the products at regular price p which is a decision variable. In the second stage, the manufacturer sells the products at lower price s which is an exogenous variable and decided by the market. The price in second stage is still greater than the cost of the product, i.e. $s > c$. The initial unit carbon emissions of the product is e_0. The manufacture can invest in green technology to reduce unit carbon emissions. The unit carbon emissions after green technology investment is $e(e < e_0)$. Referring to the expression of [16], the paper assumes the green technology investment function, denoted by $I(e)$, is:

$$I(e) = t(e_0 - e)^2, \tag{1}$$

where t represents the investment coefficient. So the manufacturer's decision variables is regular sales price p and unit carbon emissions e.

This paper characterizes customers as heterogeneous low carbon customers. The heterogeneity behavior in two aspects. On the one hand, the customers include myopic and strategic types. On the other hand, the customers have different valuation to the product. The customers can also observe the unit carbon emissions during the sales period and take it into consideration to make a buying decision. The paper assumes that the customer size are fixed and denoted by N. The ratio of strategic customers is θ and the ratio of myopic customers is $1-\theta$. The paper assumes that each customer purchases a product at most. The valuation of the two type customers (denoted by v) is follows uniform distribution in $[0, v^f]$. Strategic customers discount future consumption at δ, $(0 < \delta < 1)$. By delaying a purchase until the second stage, Strategic customers lose out on some consumptive value, and hence their future utility is reduced to reflect this loss [3]. δ can be see as patience of the strategic customer and higher δ implies more patient customers. The unit cost is c. The shortage cost and salvage value is not take into consideration.

In addition, $v^f > p + \eta e_0 > s + \eta e_0$, which indicates that the customers have positive margin when they buy the products. In order to express convenient, set $A = \frac{1-\delta+\theta\delta}{1-\delta}$.

3 The Base Model

3.1 Analysis of Decision Making Process

According to traditional study, the decision condition for myopic customers is $v - p > 0$. This paper extends it as $v - p - \eta e_0 > 0$ to depict the carbon sensitive characteristics of customers, where η represents the carbon sensitive degree. The extension of strategic customer decision rules is similar, that is $v - p - \eta e_0 > 0$ and $v - p - \eta e_0 > \delta(v - s - \eta e_0)$.

(1) The first stage

At the first stage, the manufacturer set p. Then the two type customers will response to these decisions according to themselves rules.

As to myopic customers. According to the decisions rules, myopic customers will buy products as long as $v-p-\eta e_0>0$, Otherwise, they will not buy products and leave the market. So the expected number of myopic customers who take the purchase action in the first stage is $(1-\theta)N\int_{p+\eta e_0}^{v^f} f(v)\,dv=(1-\theta)N\frac{v^f-p-\eta e_0}{v^f}$.

As to strategic customers. According to the decisions rules, strategic customers will buy products only if:

$$v-p-\eta e_0>0,\ v-p-\eta e_0>\delta(v-s-\eta e_0).$$

Then

$$0<\frac{p-\delta s+(1-\delta)\eta e_0}{1-\delta}<v<v^f.$$

So the expected number of strategic customers who take the purchase action in the first stage is:

$$\theta N\int_{\frac{p-\delta s+(1-\delta)\eta e_0}{1-\delta}}^{v^f} f(v)\,dv=\theta N\frac{v^f-\frac{p-\delta s}{1-\delta}-\eta e_0}{v^f}.$$

To sum up, the expected number of customers who take the purchase action in the first stage is $(1-\theta)N\int_{p+\eta e_0}^{v^f} f(v)\,dv=(1-\theta)N\frac{v^f-p-\eta e_0}{v^f}$.

(2) The second stage

At the second stage, the sales price reduces to s. Then the rest of the two type customers will response to this price according to themselves rules.

As to myopic customers. The expected number who take the purchase action in the second stage is $(1-\theta)N\int_{s+\eta e_0}^{p+\eta e_0} f(v)\,dv=(1-\theta)N\frac{p-s}{v^f}$.

As to strategic customers. The product will exit the market after the second stage. Waiting does not make sense for strategic customers now. As long as the valuation is greater than $s+\eta e$, the strategic customers will buy products. Therefore, the expected number of strategic customers who take the purchase action in the second stage is: $\theta N\int_{s+\eta e_0}^{\frac{p-\delta s+(1-\delta)\eta e_0}{1-\delta}} f(v)\,dv=\theta N\frac{\frac{p-s}{1-\delta}+\eta e_0}{v^f}$.

Based on the above analysis, the number of customers who purchase products can be obtained and listed in Table 1.

Table 1. The number of customers who purchased the product

	First stage	Second stage
Myopic customers	$(1-\theta)N\frac{v^f-p-\eta e_0}{v^f}$	$(1-\theta)N\frac{p-s}{v^f}$
Strategic customers	$\theta N\frac{v^f-\frac{p-\delta s}{1-\delta}-\eta e_0}{v^f}$	$\theta N\frac{\frac{p-s}{1-\delta}+\eta e_0}{v^f}$
Total	$\frac{N}{v^f}\left[v^f-(1-\theta)p-\frac{\theta(p-\delta s)}{1-\delta}-\eta e_0\right]$	$\frac{N}{v^f}\left[\frac{(1-\delta+\theta\delta)(p-s)}{1-\delta}+\eta e_0\right]$

3.2 Pricing Model Without Green Technology Investment

The manufacturers expected profit function without green technology investment, denoted by $\pi_n(p)$, is:

$$\begin{aligned}\pi_n(p) =& \frac{N}{v^f}\left[v^f-(1-\theta)\,p-\frac{\theta\,(p-\delta s)}{1-\delta}-\eta e_0\right](p-c)\\ &+\frac{N}{v^f}\left[\frac{(1-\delta+\theta\delta)\,(p-s)}{1-\delta}+\eta e_0\right](s-c)\,.\end{aligned} \tag{2}$$

Lemma 1. *$\pi_n(p)$ is a concave function of p.*

Proof. According to formula (2), I have $\frac{d\pi_n(p)}{dp}=\frac{N}{v^f}\left[v^f-\frac{2(1-\delta+\theta\delta)}{1-\delta}p-\eta e_0+\frac{1-\delta+2\theta\delta}{1-\delta}s\right]$, $\frac{d^2\pi_n(p)}{dp^2}=-\frac{N}{v^f}\frac{2(1-\delta+\theta\delta)}{1-\delta}<0$. So I can obtain that $\pi_n(p)$ is a concave function of p. This completes the proof. □

Proposition 1. *There exists an unique optimal pricing policy p_n, which maximizes $\pi_n(p)$, where $p_n=\frac{v^f}{2A}-\frac{\eta e_0}{2A}+\left(1-\frac{1}{2A}\right)s$.*

Proof. Set $\frac{d\pi_n(p)}{dp}=0$, I have $v^f-\frac{2(1-\delta+\theta\delta)}{1-\delta}p-\eta e_0+\frac{1-\delta+2\theta\delta}{1-\delta}s=0$. Assume $A=\frac{1-\delta+\theta\delta}{1-\delta}$, I can derive the optimal pricing policy $p_n=\frac{v^f}{2A}-\frac{\eta e_0}{2A}+\left(1-\frac{1}{2A}\right)s$ by solving the equation. This completes the proof. □

Proposition 1 shows that the manufacturer's optimal pricing policy without green technology investment is unique exist.

Proposition 2. *p_n is decreasing in θ.*

Proof. $\frac{dp_n}{d\theta}=\frac{dp_n}{dA}\frac{dA}{d\theta}=-\frac{v^f-\eta e_0-s}{2A^2}\frac{\delta}{1-\delta}$. I have $v^f-\eta e_0-s>0$, $1-\delta>0$. Then I can obtain $\frac{dp_n}{d\theta}<0$, i.e. p_n is decreasing in θ. This completes the proof. □

This conclusion is intuitive. θ represents the ratio of strategic customer. When the number of strategic customers is increasing, in order to enable more strategic customers buying at regular price, the manufacture will reduce the regular price in first stage.

Proposition 3. *p_n is decreasing in δ.*

Proof. $\frac{dp_n}{d\delta}=\frac{dp_n}{dA}\frac{dA}{d\delta}=-\frac{v^f-\eta e_0-s}{2A^2}\frac{\theta}{(1-\delta)^2}$. I have $v^f-\eta e_0-s>0$. Then I can obtain $\frac{dp_n}{d\delta}<0$, i.e. p_n is decreasing in δ. This completes the proof. □

Proposition 3 shows that the optimal pricing of the manufacturer decreases in consumer patience (i.e. δ). More patience means loss out less consumptive value and more willing to buy the product in second stage for the strategic customers. In order to stimulate strategic customers buy early, the manufacturer have to reduce the price.

Proposition 4. *p_n is decreasing in η.*

Proof. $\frac{dp_n}{d\eta} = -\frac{e_0}{2A^2} < 0$. So $\frac{dp_n}{d\eta} < 0$, i.e. p_n is decreasing in η. This completes the proof. □

Proposition 4 indicates that the optimal pricing of the manufacturer decreasing in the carbon sensitive degree (i.e. η). Without green technology investment, the unit carbon emissions is fixed and higher η means the customers less willing to buy the products. In order to stimulate more customers to buy products, the manufacturer will reduce the optimal price in first stage.

4 Pricing and Greening Model with Green Technology Investment

4.1 The Optimal Policy

The analysis of decision making process is similar to the base model and the difference is that the unit carbon emissions is no longer constant but a decision variable of the manufacturer. According to the analysis in Sect. 2, I can obtain the manufacturer's expected profit function, denoted by $\pi_g(p,e)$, with green technology investment.

$$\begin{aligned}\pi_g(p,e) =& \frac{N}{v^f}\left[v^f - (1-\theta)p - \frac{\theta(p-\delta s)}{1-\delta} - \eta e\right](p-c) \\ &+ \frac{N}{v^f}\left[\frac{(1-\delta+\theta\delta)(p-s)}{1-\delta} + \eta e\right](s-c) + t(e_0-e)^2.\end{aligned} \tag{3}$$

Lemma 2. *If $4tA - \eta^2 > 0$, $\pi_g(p,e)$ is a joint concave function of p and e.*

Proof. I can obtain the partial derivatives of profit function $\pi_g(p,e)$ with respect to p and e:

$$\frac{\partial \pi_g(p,e)}{\partial p} = \frac{N}{v^f}\left[v^f - \frac{2(1-\delta+\theta\delta)}{1-\delta}p - \eta e + \frac{1-\delta+2\theta\delta}{1-\delta}s\right], \tag{4}$$

$$\begin{aligned}\frac{\partial^2 \pi_g(p,e)}{\partial p^2} &= -\frac{2(1-\delta+\theta\delta)}{1-\delta} < 0, \\ \frac{\partial \pi_g(p,e)}{\partial e} &= -\frac{N}{v^f}(p-s) + 2t(e_0-e), \\ \frac{\partial^2 \pi_g(p,e)}{\partial e^2} &= -2t < 0, \\ \frac{\partial^2 \pi_g(p,e)}{\partial p \partial e} &= \frac{\partial^2 \pi_g(p,e)}{\partial e \partial p} = -\eta.\end{aligned} \tag{5}$$

I have $4tA - \eta^2 > 0$, then I can show that

$$\begin{vmatrix} \frac{\partial^2 \pi_g(p,e)}{\partial p^2} & \frac{\partial^2 \pi_g(p,e)}{\partial p \partial e} \\ \frac{\partial^2 \pi_g(p,e)}{\partial e \partial p} & \frac{\partial^2 \pi_g(p,e)}{\partial e^2} \end{vmatrix} = 4tA - \eta^2 > 0.$$

So $\pi_g(p,e)$ is a concave function of p and e. This complete the proof. □

Proposition 5. *The optimal pricing and greening policy (denoted by p_g and e_g) of the manufacturer are:*

$$p_g = \frac{2tv^f\left[v^f - e_0\eta + (2A-1)\,s\right] - N\eta s}{4Atv^f - N\eta},$$
$$e_g = \frac{4Ae_0tv^f - N\left(v^f - s\right)}{4Atv^f - N\eta}.$$

Proof. Let $\frac{\partial \pi_g(p,e)}{\partial p} = 0$ and $\frac{\partial \pi_g(p,e)}{\partial e} = 0$, from equation (4) and (5), I get $p_g = \frac{2tv^f[(2A-1)s - e_0\eta + 1] - N\eta s}{4Atv^f - N\eta}$ and $e_g = \frac{4Ae_0tv^f - N\left(v^f - s\right)}{4Atv^f - N\eta}$. This complete the proof. □

Proposition 5 show that the optimal pricing and greening policy for the manufacturer with green technology investment exist and are unique.

Proposition 6. *p_g is decreasing in θ; e_g is increasing in θ.*

Proof. From Proposition 5, I get the value of p_g and e_g, then I can show

$$\frac{dp_g}{d\theta} = \frac{dp_g}{dA}\frac{dA}{d\theta} = -\frac{8t^2{v^f}^2\left(v^f - \eta e_0 - s\right)}{\left(4Atv^f - N\eta\right)^2}\frac{\delta}{1-\delta},$$
$$\frac{de_g}{d\theta} = \frac{dp_g}{dA}\frac{dA}{d\theta} = \frac{4Ntv^f\left(v^f - \eta e_0 - s\right)}{\left(4Atv^f - N\eta\right)^2}\frac{\delta}{1-\delta}.$$

I have $v^f - \eta e_0 - s > 0$. Then I can obtain $\frac{dp_g}{d\theta} < 0$ and $\frac{de_g}{d\theta} > 0$. Therefore, p_g is decreasing in θ and e_g is increasing in θ. This complete the proof. □

Proposition 6 shows that the higher the ratio of strategic customers, the lower the manufacturer's optimal price. This conclusion is in line with the scenario without green technology investment.

Proposition 7. *p_g is decreasing in δ; e_g is increasing in δ.*

Proof. I can obtain

$$\frac{dp_g}{d\delta} = \frac{dp_g}{dA}\frac{dA}{d\delta} = -\frac{8t^2{v^f}^2\left(v^f - \eta e_0 - s\right)}{\left(4Atv^f - N\eta\right)^2}\frac{\theta}{(1-\delta)^2},$$
$$\frac{de_g}{d\theta} = \frac{dp_g}{dA}\frac{dA}{d\theta} = \frac{4Ntv^f\left(v^f - \eta e_0 - s\right)}{\left(4Atv^f - N\eta\right)^2}\frac{\delta}{1-\delta}.$$

I have $v^f - \eta e_0 - s > 0$. Then I can obtain $\frac{dp_g}{d\delta} < 0$ and $\frac{de_g}{d\delta} > 0$. Therefore I get p_g is decreasing in δ and e_g is increasing in δ. This complete the proof. □

Proposition 7 indicates that more patience the strategic customer have, lower the optimal pricing with green technology investment is and higher the optimal unit carbon emissions is with green technology investment. As to the impact of δ on the optimal unit carbon emissions, I find that more "strategic" customers will lead to less investment that the manufacturer do in carbon reduction.

Proposition 8. *If* $N > \frac{4Atv^f}{\eta}$, p_g *and* e_g *are decreasing in* η.

Proof. $N > \frac{4Atv^f}{\eta}$, then $4Atv^f < N\eta$. Recall $v^f - \eta e_0 - s > 0$, so

$$\frac{dp_g}{d\eta} = \frac{2Ntv^f\left(v^f - s\right) - 8Ae_0t^2v^{f^2}}{\left(4Atv^f - N\eta\right)^2} > \frac{2Ntv^f\left(v^f - \eta e_0 - s\right)}{\left(4Atv^f - N\eta\right)^2} > 0,$$

$$\frac{de_g}{d\eta} = -\frac{N^2\left(v^f - s\right) - 4NAe_0tv^f}{\left(4Atv^f - N\eta\right)^2} < -\frac{N^2\left(v^f - \eta e_0 - s\right)}{\left(4Atv^f - N\eta\right)^2} < 0.$$

□

Proposition 8 shows that with the addition of carbon sensitive degree (i.e. η), the manufacturer will improve the regular sale price and invest more in green technology investment to reduce carbon reduction when the customer number exceed a certain threshold.

4.2 Effect of Green Technology Investment

In order to analyze the effect of green technology investment on the manufacture's optimal pricing, I compare the manufacturer's optimal prices with and without green technology investment.

Proposition 9. *If* $N > \frac{4Atv^f}{\eta}$, $p_g < p_n$.

Proof. $p_g - p_n = \frac{2tv^f\left[v^f - e_0\eta + (2A-1)s\right] - N\eta s}{4Atv^f - N\eta} - \frac{v^f}{2A} - \frac{\eta e_0}{2A} + \left(1 - \frac{1}{2A}\right)s = \frac{N\eta\left(v^f - \eta e_0 - s\right)}{2A\left(4Atv^f - N\eta\right)}$. If $N > \frac{4Atv^f}{\eta}$, i.e. $4Atv^f - N\eta < 0$, then $p_g - p_n < 0$, i.e. $p_g < p_n$. This complete the proof. □

Proposition 9 indicates that the optimal price of the manufacturer with green technology investment is lower than that of without green technology investment when the number of customers exceed a certain threshold. Considering green technology investment, the manufacturer reduces the unit carbon emissions and then get the room for price reduction. In order to maximize the expected profit, the manufacturer need to reduce the optimal price to enable more customers to buy the products.

5 Conclusion and Future Research

This paper investigates the pricing and greening policy of a manufacturer and analyze the impact of green technology investment on the optimal price. The manufacture production a product and sells them to the mix myopic and strategic customers. Customers will be affected by unit carbon emissions of the product when making a purchase decision. The manufacture can invest in green technology to reduce unit carbon emissions. To the best of our knowledge, this is the first study on the model incorporate mix myopic and strategic customers and green technology investment together. This article provides several interesting observations.

(1) There is an unique optimal pricing policy for the manufacturer without green technology investment and there is also an unique optimal pricing and greening policy for the manufacture with green technology investment. Therefore, by deriving the optimal policies in different scenarios, the manufacturer can behave appropriately based on the findings to maximize its expected profit.
(2) I find that the optimal price without green technology investment is decreasing in the ratio of strategic customers, patience of strategic customers and carbon sensitive degree. This conclusion means that the presence of strategic customers is detrimental to the manufacturer and the negative impact increases as the ratio and patience of strategic customers increases.
(3) I find that the optimal price with green technology investment is also decreasing in the ratio and patience of strategic customers and carbon sensitive degree; the optimal unit carbon emissions is increasing in the ratio and patience of strategic customers but decreasing in carbon sensitive degree.
(4) The optimal price with green technology investment is lower than that without green technology investment when the number of customers exceed a certain threshold. This conclusion is interesting. The cost of the product is increasing when the manufacturer invest in green technology. According to the experience, the optimal price should be improved. However, the optimal pricing of the manufacturer not increased but reduced instead. This conclusion can guide firms make decisions scientific.

The paper formulates the model by considering only one manufacturer. But supply chain is more common in real-life. Extending this assumption will have a knock-on effect on supply chain decisions. So one key research direction is to consider a supply chain scenario. The manufacturer decide the wholesale price and green technology investment and the retailer decide selling price and order quantity. Second, this paper assume that the manufacturer only produces one product. While this simplified setting provides some interesting insights, we have to acknowledge that usually two or more products is produced by a manufacturer. These products will substitute or complement each other. So to assume that the manufacturer produces two or more products, which have different unit carbon emissions in production, is one important extension of this work.

Acknowledgements. This research is partially supported by National Natural Science Foundation of China (No. 71602134) and Research Interest Training Program of Sichuan Agricultural University (No. 2016133).

References

1. Aviv Y, Pazgal A (2008) Optimal pricing of seasonal products in the presence of forward looking consumers. Manufacturing Serv Oper Manag 3(10):339–359
2. Bansal S, Gangopadhyay S (2003) Tax/subsidy policies in the presence of environmentally aware consumers. J Environ Econ Manag 45(2):333–355
3. Cachon GP, Swinney R (2011) The value of fast fashion: quick response, enhanced design, and strategic consumer behavior. Manag Sci 4(57):778–795

4. Du J, Zhang J, Hua G (2015) Pricing and inventory management in the presence of strategic customers with risk preference and decreasing value. Int J Prod Econ 164:160–166
5. IPCC (2007) Cliamte change 2007 synthesis report. IPCC pp 1–13
6. Jerath K, Netessine S, Veeraraghavan SK (2010) Revenue management with strategic customers: Last-minute selling and opaque selling. Manage Sci 3(56):430–448
7. Jiang W, Chen X (2012) Manufacturers production and pricing policy with carbon constraint and strategic customer behavior. Adv Inf Sci Serv Sci 23(4):231–238
8. Jiang W, Chen X (2015) Supply chain decisions and coordination with strategic customer behavior under cap-and-trade policy. Control Decis 31(3):477–485
9. Jiang W, Chen X (2016) Optimal strategies for low carbon supply chain with strategic customer behavior and green technology investment. Discrete Dynamics Nat Soc, 1–13
10. Jiang W, Chen X (2016) Optimal strategies for manufacturer with strategic customer behavior under carbon emissions-sensitive random demand. Ind Manag Data Syst, 4(116):759–776
11. Klingelhö FHE (2009) Investments in eop-technologies and emissions trading-results from a linear programming approach and sensitivity analysis. Eur J Oper Res 1(196):370–383
12. Lai G, Debo LG, Sycara K (2010) Buy now and match later: Impact of posterior price matching on profit with strategic consumers. Manufacturing Serv Oper Manag 1(1):33–55
13. Prasad A, Venkatesh R, Mahajan V (2015) Product bundling or reserved product pricing? price discrimination with myopic and strategic consumers. Int J Res Mark 32(1):1–8
14. Su X (2007) Intertemporal pricing with strategic customer behavior. Manage Sci 5(53):726–741
15. Xia D, Chen B, Zheng Z (2015) Relationships among circumstance pressure, green technology selection and firm performance. J Clean Prod 106:487–496
16. Yalabik B, Fairchild RJ (2011) Customer, regulatory, and competitive pressure as drivers of environmental innovation. Int J Prod Econ 2(131):519–527
17. Zhao J, Hobbs BF, Pang JS (2010) Long-run equilibrium modeling of emissions allowance allocation systems in electric power markets. Oper Res 3(58):529–548

Evolutionary Game Analysis of the Reservoir Immigrants' Joint Venture

Xiaofeng Liu(✉), Tingting Song, and Yanhua Liu

The Management Department in Northwest University for Nationalities, Gansu, People's Republic of China
liuxiaofengmdqk@126.com

Abstract. This paper combined the previous results with the actual situation of the immigrant area, proposed the main factors which would influence the joint venture, such as technology level, the familiar degree of both sides of cooperation, the costs and benefits of cooperation and coordination. To establish the evolutionary game model of immigrant joint venture, Finding the influence factors and the "free-rider" punish coefficient, the investment allocation coefficient, the income distribution coefficient how to affect the direction of the joint venture for immigrants according to payoff matrix and the Jacobi matrix.

Keywords: The reservoir immigrants · The joint venture · Evolutionary game

1 Introduction

With the time increase of reservoir immigrants living in the settlement, the possibility of its entrepreneurship. Many immigrants chose to entrepreneurship on some basis of technical in Yongjing county in Linxia Hui Autonomous Prefecture, Gansu province, such as tourism, fisheries and greenhouse cultivation. It still exists social employment problem now, the government vigorously encouraging entrepreneurship. There are many immigrants started the business in Yongjing county, but to reduce risk, choosing partners to joint venture will reduce the risk of entrepreneurship. Therefore, it is very necessary to study on the joint venture.

Yang [6] used the binary Logistic regression method to study the influence factors of immigrant entrepreneurs. Yang [7] re-examined the "bounded rationality" of venture capitalists. Su [3] argued that the policy of encouraging cooperation that government put forward to has not work, suggesting that the government should intensify the policy support and increasing investment. Chou [2] suggested when cooperating, both firms need to contribute sufficient and complementary efforts when they choose to cooperate partner. When entrepreneurs choosing a partner, relation network is the key variables to affect entrepreneurial decision [1]. Szolnoki [4] thouhgt that a carefully chosen threshold to establish a joint venture could efficiently improve the cooperation level. Wu and Wang [5] thought that the harder the joint venture is, the higher level of cooperation will be.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_106

2 Evolutionary Game Model

Evolutionary game theory is the result of the evolution, which is a kind of dynamic equilibrium.

2.1 The Assumption of Model Variables

In this paper, the mode of entrepreneurship is individual, cooperation means people to do a project together. Assuming that the probability of individual A choose to cooperation is x, uncooperative probability is $1-x$, and the probability of individual B choose to cooperation is y, uncooperative probability is $1-y$. The rest of the variables are as follows: The technical level T_i, Resource utilization ability α_i, The cost of coordination C, the prior information Q_i, The credibility of the prior information β_i, the profit R, the free-rider penalty coefficient d. To make the following *Hypothesis* about the problem of the joint venture:

(1) This study involved two entrepreneurs, assuming that there are two individuals A and B, and the two individuals are in the same industry, they have their own worker sources and experience. In their cooperative game has two choices: cooperation and noncooperation.
(2) When the cooperative game beginning, both sides will take the degree of familiar with each other into account, a person's credit affects whether other people can choose to believe him. Both sides have some prior information about each other, which is the understanding of each other, but there is a credibility in prior information, and some of the information may be wrong. So it is important to put the prior information reliability coefficient as β_i. That is to say, A has the prior information with B, which is Q_B, The credibility of the information is β_A.
(3) Both the game sides have a certain technology. What's more, technical degree is one of the factors that both sides of game will consider in choosing cooperation. Assuming that technical level is T_i, there also is exists a certain degree of resource utilization while working together, assuming that resource utilization coefficient is α_i.
(4) Assuming that the investment apportion coefficient are a and $1-a$, which can get the income distribution coefficient b and $1-b$.
(5) When there is the phenomenon of "free-rider", it will have some punishment d, the punish coefficient is established on the basis of the technical level.

2.2 Evolutionary Game Model

According to the hypothesis and model variables, we can get the payoff matrix of A and B, as shown in Table 1.

In order to facilitate discussion, assuming that the value of x and y is invariable. To calculate the expectations of A and B in the case of cooperation and noncooperation respectively, and then get replicated dynamic equation of A and B:

Table 1. The payoff matrix

		B	
		Cooperation y	Noncooperation $1-y$
A	Cooperation x	$bR+\beta_A Q_B+\alpha_A T_B-aC$, $(1-b)R+\beta_B Q_A+\alpha_B T_A-(1-a)C$	$bR-aC-T_A$ $(1-b)R+\alpha_B T_A-(1-a)$ $C-dT_A$
	Noncooperation $1-x$	$bR+\alpha_A T_B-aC-dT_B$, $(1-b)R-(1-a)C-T_B$	0,0

$$\begin{aligned}
\text{A:}\quad F(x) = \frac{dx}{dt} &= x(U_{A1}-U_A) = x(1-x)[y(\beta_A Q_B + dT_B - bR + aC + T_A) \\
&\quad + (bR - aC - T_A)], \\
\text{B:}\quad F(y) = \frac{dy}{dt} &= y(1-y)\{x[(\beta_B Q_A + (1-a)C + dT_A + T_B - (1-b)R] \\
&\quad + (1-\mathrm{b})R - (1-a)C - T_B\}
\end{aligned}$$

To obtain the evolutionary stable strategy of the game, First of all, we need to get the stable point of replicated dynamic equation, namely it need to meet such conditions:

No. 1 $F(\mathrm{x}) = 0$; No. 2 Evolutionary stable strategy needs to have some anti-interference. For $\begin{cases} F(x)=0 \\ F(y)=0 \end{cases}$, we can get five stable points, O$(0,0)$, A$(0,1)$, B$(1,0)$, C$(1,1)$, D$(\mathrm{X_D},\mathrm{Y_D})$.

$$\begin{aligned}
X_D = x^* &= \frac{-(1-b)R + (1-a)C + T_B}{\beta_B Q_A + dT_A - (1-b)R + (1-a)C + T_B}, \\
Y_D = y^* &= \frac{-bR + aC + T_A}{\beta_A Q_B + dT_B - bR + aC + T_A}.
\end{aligned}$$

Assuming that $M = \beta_A Q_B + dT_B$ and $W = bR - aC - T_A$ in $F(x)$; $E = \beta_B Q_A + dT_A$ and $F = (1-b)R - (1-a)C - T_B$ in $F(y)$.

According to the methods put forward by Friedman, the dynamic equilibrium of participants' activities can be got according to the part of the system stability of the Jacobi matrix. We can get the Jacobi matrix based on the above. $J = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$, Among them,

$$\begin{aligned}
X_{11} &= (1-2x)[y(M-W)+W], \\
X_{12} &= \mathrm{x}(1-x)(M-W), \\
X_{21} &= y(1-y)(E-F), \\
X_{22} &= (1-2y)[x(E-F)+F].
\end{aligned}$$

We use tr$J$ and det J to express the trace of matrices and determinant. When the stable points meet the conditions: tr$J > 0$ and det $J > 0$, The stable point is the evolutionary stable strategy of the system (ESS), the value of the Jacobi matrix and determinants are as follow:

$$tr J = X_{11} + X_{22},$$
$$\det J = X_{11}X_{22} - X_{12}X_{21}.$$

Based on the contents mentioned above we can get the local stability of equilibrium under different parameters:

(1) When $W < 0$, $F < 0$, $M < 0$, $E < 0$, the evolutionary stable strategy of the system is $\mathrm{O}(0,0)$.
(2) When $M < 0$, $E < 0$, $W > 0$, $F > 0$, the evolutionary stable strategy of the system is $\mathrm{A}(0,1)$, $\mathrm{B}(1,0)$.
(3) When $W > 0$, $E > 0$, $M > 0$, $F > 0$, the evolutionary stable strategy of the system is $\mathrm{C}(1,1)$.

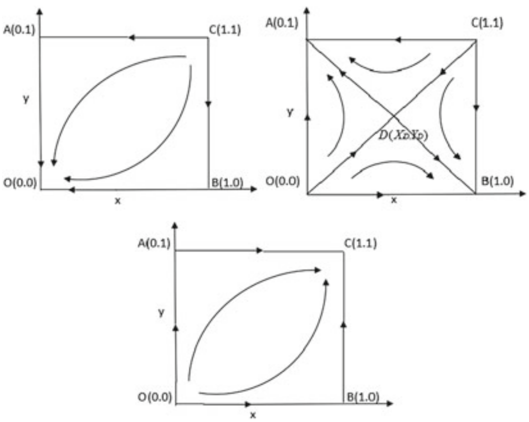


Fig. 1. System dynamic phase diagram of three different conditions

2.3 The Influence Factors Analysis of Evolution Path

Because of the coefficient set in the evolutionary game model that may influence the final result, the probability of system dynamic evolution reached to $\mathrm{O}(0,0)$, $\mathrm{A}(0,1)$, $\mathrm{B}(1,0)$, $\mathrm{C}(1,1)$. Probability of cooperation can be determined according to the area of the quadrilateral and quadrilateral. Therefore, for discussing the effect of various parameters on the dynamic evolution of the payoff matrix can be converted into the effects of various parameters on the area of the quadrilateral. The area of can be expressed as:

$$S_{ADBC} = \frac{(1-y^*)}{2} + \frac{(1-x^*)}{2} = \frac{E}{2(E-F)} + \frac{M}{2(M-W)},$$

(1) Request the first partial derivative and second derivative of S_{ADBC} about a, we can get $\frac{\partial^2 S_{ADBC}}{\partial a^2} = \frac{(E-F)C^2E}{(E-F)^4} + \frac{(M-W)C^2M}{(M-W)^4} > 0$.
  - If $E > M$, that is to say when investment apportion coefficient get greater, the both sides of game tend to be non-cooperative.
  - If $E < M$, that is to say when investment apportion coefficient get greater, the both sides of game tend to be cooperative.

(2) Request the first partial derivative and second derivative of S_{ADBC} about b, we can get $\frac{\partial^2 S_{ADBC}}{\partial b^2} = \frac{(E-F)R^2E}{(E-F)^4} + \frac{(M-W)M^2R}{(M-W)^4}$.
  - If $E > M$, that is to say when income distribution coefficient get greater, the both sides of game tend to be cooperative.
  - If $E < M$, that is to say when get income distribution coefficient greater, the both sides of game tend to be non-cooperative.

(3) Request the first partial derivative S_{ADBC} about C, then can get $\frac{\partial S_{\mathrm{ADBC}}}{\partial C} = \frac{1}{2}\left[\frac{(a-1)E}{(E-F)^2} + \frac{-aM}{(M-W)^2}\right] > 0$, that is to say when the coordination cost get greater, the evolutionary games tend to be non-cooperative.

(4) Request the first partial derivative S_{ADBC} about d, when the penalty coefficient get greater, the evolutionary game tends to be the cooperative.

(5) Request the first partial derivative S_{ADBC} about β_A and β_B, we can conclude that when the familiar degree between A and B is more higher, the more tendency for them to cooperate.

(6) Request the first partial derivative S_{ADBC} about α_{i}, the result is 0, so the ability to absorb and utilize resources in the cooperation is not the main factors.

2.4 The Analysis of Model Results

Based on the analysis above, we can conclude that:

(1) When the value of a and b equal to 0.5, it will reach the steady state, namely when both the investment allocation coefficient and income distribution coefficient are 0.5, the probability of cooperation is the largest.
(2) The less the coordination costs are, the more the probability of cooperation is.
(3) The bigger the "free-rider" punish coefficient is, the bigger the probability of cooperation is.
(4) The higher the credibility of a prior information is, the bigger the probability of cooperation is.
(5) The ability to absorb and utilize resources in the cooperation is not the main factors.

3 The Analysis of an Empirical Example

Through investigation and interview in Yongjing county in Linxia Hui Autonomous Prefecture, Gansu province, Some of immigrants are first to migrate, some are the second. Expert scoring method is used to evaluate several

Table 2. The value of the various factors

The symbol	The value
$T_{\rm i}$	$T_A = 5, T_B = 6$
$\alpha_{\rm i}$	$\alpha_A = 0.6, \alpha_B = 0.7$
C	500
$Q_{\rm i}$	$Q_A = 2, Q_B = 3$
$\beta_{\rm i}$	$\beta_A = 0.2, \beta_B = 0.1$
R	5000
d	0.3

related aspects on some individuals according to the assumption. We found out two individual households at same level, The value of the various factors are as follows (Table 2):

According to the conclusion of evolutionary game model, assuming that $a = b = \frac{1}{2}$, and setting the value of the different factors according to the hypothesis, and then we can get the payoff matrix with different value:

Table 3. The No. 1 payoff matrix

		B	
		Cooperation y	Noncooperation $1 - y$
A	Cooperation x	2254.2, 2253.7	2245, 2252
	Noncooperation $1 - x$	2251.8, 2244	0, 0

It is the best choice for A and B to cooperate with each other. When $T_A = 1, T_B = 9$, other value isn't changed, the payoff matrix is as follow:

Table 4. The No. 2 payoff matrix

		B	
		Cooperation y	Noncooperation $1 - y$
A	Cooperation x	2256, 2250.9	2249, 2250.4
	Noncooperation $1 - x$	2252.7, 2241	0, 0

Comparing Table 3 with Table 4, it is not difficult to find that, in Table 4, the benefits is similar for B to choose cooperation or not, so the smaller technology differences between both sides of the game is, the more tendency to cooperation. It's easy to prove the other results from model analysis with this methods.

4 Conclusion

According to the example of Yongjing county in Linxia Hui Autonomous Prefecture, Gansu province, we can get the main factors which would influence the joint venture in this area, such as technology level, the familiar degree of both sides of cooperation, the costs and benefits of cooperation and coordination. we can also get the direction of every factors and the “free-rider” punish coefficient, the investment allocation coefficient, the income distribution coefficient influence on joint venture.

(1) When the both sides of the game have a big difference on technical level, the high level of technical would not choose joint venture.
(2) The most residents in settlement came from different areas, and some of them experienced two or more times of migration. When they familiar with each other, they will have a sense of security, and the higher the credibility of the information is, the more possibility the choice of the joint venture will be.
(3) The higher costs and benefits of cooperation is, the greater the risk will be, and the probability of joint venture will decline.
(4) The greater the benefits of joint venture is, the greater the probability of cooperation will be.
(5) The bigger the “free-rider” punish coefficient is, the bigger the probability of cooperation is.
(6) When both the investment allocation coefficient and income distribution coefficient are 0.5, the probability of cooperation is the largest.

Acknowledgement. This research was financially supported by Ministry of Education of the People’s Republic of China Project (12YIA630080).

References

1. Ardichvili A, Cardozo R, Ray S (2003) A theory of entrepreneurial opportunity identification and development. J Bus Ventur 18(1):105–123
2. Chou PB, Bandera C, Thomas E (2017) A behavioural game theory perspective on the collaboration between innovative and entrepreneurial firms. Int J Work Innovation 2(1):6–31
3. Su L, Yang Z et al (2013) The influence factors of green technology cooperation between enterprises, based on the perspective of supply chain. The Population of China Resources and environment, pp 149–154 (in Chinese)
4. Szolnoki A, Chen X (2016) Cooperation driven by success-driven group formation. Phys Rev E 94(4):042–311
5. Wu T, Fu F, Zhang Y et al (2013) The increased risk of joint venture promotes social cooperation. PLoS One 8(6):1–10
6. Yang X, Wang C, Xiong Y (2015) The influence factors study of the three gorges reservoir immigrant entrepreneurship decision. The Rural Economy pp 120–124 (in Chinese)
7. Yang Y (2016) Based on evolutionary game of joint venture formation mechanism research. J Technol Econ Manag Res 12:20–24 (in Chinese)

Solid Waste Management in the Republic of Moldova

Gheorghe Duca(✉) and Aliona Mereuţa

Academy of Sciences of Moldova, Chisinau, Republic of Moldova
malvina.condratiuc@mail.ru

Abstract. Waste management refers to all operations implemented for the proper management of solid waste: methods of collecting, recycling and recovery, processing, treatment, storage, transportation, analysis and monitoring. The main topics in this paper are the following: institutional framework, national legislative and regulatory frameworks of waste management; Integrated management of municipal waste; Hazardous waste management; Secondary wine product management; Management of plastic waste; Infectious medical waste management. Thus, ASM in collaboration with a private company implemented two technologies for complex treatment of solid waste. The following goals were currently reached in RM: a sector of complex full processing of plastic waste was created at industrial scale, a pilot line was put into operation, for sorting, processing and obtaining of new products from secondary polymer raw materials and an authorized point was created for managing infectious waste resulting from medical activities in medical institutions and hospitals, by collecting, sorting, packing, transporting and neutralizing these wastes to prevent environmental contamination. The neutralization of infectious medical waste is performed by vacuum sterilization method using the STERISHRED 250, which is the most modern method of solving local problems of infectious medical waste processing without incineration.

Keywords: Waste management · Secondary wine product management · Integrated management of municipal waste · Management of plastic waste · Infectious medical waste management · Republic of Moldova

1 Introduction

Transforming waste management and recycling practices in Republic of Moldova is one of the most important challenges we face of the next decade. Inadequate waste management affects our communities, threatens our environment, and contributes to the global emissions of greenhouse gases. Thus, solid waste management and recycling is a local, national and international priority [1,3,5,6,9,10,12]. Government will establish the legal and institutional framework necessary to support the gradual alignment of our waste management prac-

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_107

tices to those of the European Union: Government Action Programme (2016–2018); National Strategy on Waste Management (2013–2023): Government Decision no. 428 of 10 April 2013; Strategy on Environment Protection: Government Decision no. 301 of 24 April 2014; National Plan on Implementation of the Association Agreement (2014–2016); Association Agreement between the European Union and the Republic of Moldova, ratified by Law no. 112 of 02.07.2014.

1.1 Waste Management Planning in Moldova

Municipal waste generation is influenced by many factors, the most significant being the income of the population, consumer behavior, appearance on the market of new packed products and the demographic evolution. A high level of the population's income and urbanization has resulted in the generation of large quantities of waste per capita, which is in the rural areas typically range from 0.3 to 0.4 kg/inhabitant/day, 0.9 kg/place/day or more in urban areas, according to the World Bank studies. The increasing number of supermarkets, along with increased GDP per capita has increased purchases of packaged products and hence the waste generated. Demographics also influence a waste generation, people in the urban areas producing much more waste than those in the rural areas. The most common method of treating household waste in Moldova is landfill, which often represents an important source of pollution of soil and groundwater. Annually through the sanitation services from urban localities are transported about 1144–2131 thousands m^3 of solid waste [2]. Another negative aspect of inadequate waste management is that many useful and recyclable materials are stored together with the non-recyclable thus is losing a large part of their useful potential (paper, glass, metals, plastics); being mixed, chemically and biologically contaminated, their recovery is very difficult. Effective waste management planning certainly is focused on the morphological composition of solid household waste [11]. The waste morphological analysis is imperative in recycling operations establishing and in generated waste eliminating. This data has been collected from newsletters, reports, publications, while in the national statistics such information is missing (Fig. 1).

Possible scenarios of development of the domain of reference in the Republic of Moldova were determined based on statistical data analyzes and the experts assumptions, taking into account the macro-economic indices, indicate the existing strategies for development of some sectors (industry, agriculture, etc.), and namely the Gross Domestic Product Value, demographic trends, trends of industry development, tendencies of the agricultural sector development and waste management system development. According to statistical data analysis, we can ascertain that the solid household waste tendencies increasing volumes for Chisinau city are 5–15% and in rural areas, there is a solid household waste volume increase by approximately 10%. These increases are explained by the fact that in localities, as well as in regional centers was improved the sanitation services activity. Currently, an increased attention is given to the creation of the solid household waste storage capacity and respective the annual growing trends

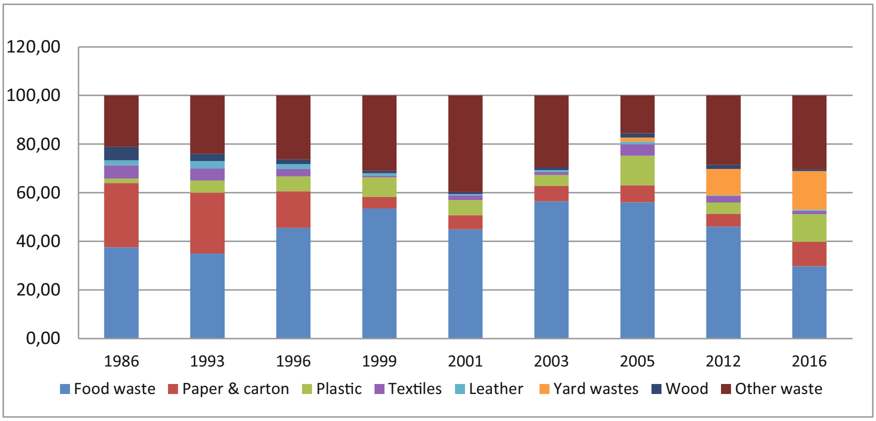


Fig. 1. Municipal waste composition

of 10% will be maintained in the 2011–2015, subsequently is expected a continued growth of 5% for both categories of localities. The solid waste generation trends are presented based on these assumptions (Fig. 2).

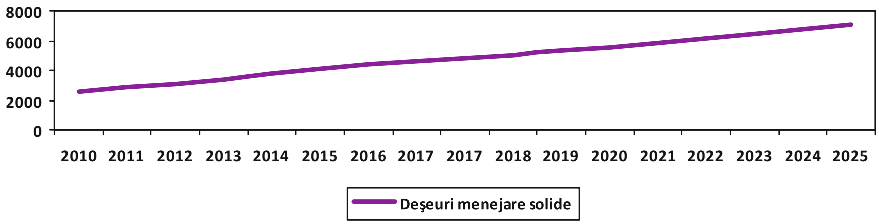


Fig. 2. The forecasts on the solid household waste generation, thousand m^3

The regional approach for waste management is essential for attracting necessary investment, as well as for the high-cost maintenance through realizing the implementation on the economical level. It is unacceptable and at the same time economically unjustified the enterprise's construction of recovery or waste disposal in every city, without taking into account the rural areas of the district, including the specific waste generated. The experience of neighboring countries shows that the object's financial sustainability of recovery and waste disposal is economically feasible when it has a territorial coverage of at least 200–300 thousand people. In this context, it is proposed the waste management regionalization by the territorial division of the country into 8 regions for waste management. The basic criteria for regional planning are the geographic position, economic development, the existence of access roads, soil, and hydrogeological conditions, the number of population, etc. (Fig. 3).

The local public authorities are expected to establish the waste management associations at a regional level, as recommended by the Ministry of Environment regarding the regional waste management planning. The role of associations is to establish and approve terms of reference for selecting the company that will manage waste in the region, as well as tariffs for the waste collection and disposal, etc.

Fig. 3. Regional waste management planning in the Republic of Moldova

Considering the country's economic development, morphological amount and composition of the generated waste, as well as pedological and geological conditions, were developed three alternatives for establishing integrated waste management system:

- Alternative 1 consists in the construction of eight solid household waste deposits for each region. The cost of implementation is approximately 293 million €.
- Alternative 2 envisions the building of 2 mechanical-biological waste treatment stations in Region 4 Chisinau and Region 7 Balti and 6 landfills for solid domestic waste for remaining regions. The cost of implementation is approximately 370 million €.
- Alternative 3 proposes the construction of 2 incinerators in Region 4 Chisinau and Region 7 Balti and 6 deposits for solid domestic waste for remaining regions. The cost of implementation is approximately 476 million €.

These options will be subject to public debate in order to select the optimal alternative from the perspective of environmental protection and cost–benefit analysis, to encourage the creation of partnerships at international, national and local level, and to attract the necessary investment that will allow the sustainable development of sector in accordance with the priority needs and at a rhythm accessible for society.

2 Advanced Treatment of Secondary Wine Products in the Republic of Moldova

For the Republic of Moldova, oenology is one of the most substantial key in the national economic sectors, thus this sector requires continuous implementation of new performances achieved in science and technology. The wine industry in the country has a potential of grape processing up to 1 million tons, located in over 100 companies. Recent years, Moldovan Wine industry attention is increasingly higher. As a result of the implementation of the State Program, the areas planted with vines have increased significantly, in the quantity of grapes processed rose as well. But it is known that about 20–25% of the processed grapes turns in are secondary wine products. If you calculate the amount of waste that result after processing 300 thousand tons of grapes, we get 60–75000 tones of products: the cake, yeast, vinasse, wastewater, etc. All these products are a source of environmental pollution and create major environmental problems. Therefore, it is important to use complex raw material (grapes), implementing technological processes with minimum waste, as well as processing products of wine in order to obtain a wide range of valuable products.

Products of vinification are an important source for obtaining special natural products with specific properties that cannot be obtained synthetically. Such products as tartaric acid, oil seed polyphones, colorings, tannins, etc. have a wide field of use in various industries.

Through the years 1980–1990, Moldova was the country that provides the raw tartar material (calcium tartrate, tirighia, calcareous sediments) for countries that produced directly tartaric acid (Ukraine, Odessa, Yerevan, Tbilisi, Italy, etc.). Prime cost of the product is quite high (100 g of tartaric acid (99.9%) used in the pharmaceutical industry costs 80–100 $, or 1 kg of natural tartaric acid used in food costs 30–40 $.

Tartaric acids are reductions that can be used in food, wine, bakery, pharmaceutical, photochemical, chemical, textile, construction, electro-technical industry, yet their production in the country is missing, although the country prevails large stocks of raw materials and annual it could produce 100 tons up to 350 tons of tartaric acid.

An effective stabilizer used in the wine industry is the met tartaric acid, which inhibits the process of submitting wine stone in wine and juice, do not change the taste of wine or juice qualities and is not toxic because it represents a modifying tartaric acid. Compared with the cold treating method, stabilization of the wines and juices with mesotartaric acid it has a much better economic effect. The local market is not assured, with this product, although it is requested on the domestic market than tartaric acid.

Grapes seed represents a precious waste for our national economy. The content of the seeds in one tone of grapes is 7% by weight. The process of obtaining seeds and their conditioning is a complicated technological process. In the years 1982–1986 in Moldova obtained to 10,000 tons seeds annually, which were transferred to processing in the Bender city (oil mill), Odessa, Armavir (Krasnodar region), Tbilisi (Georgia) and Cocand (Uzbekistan). Partly some seeds were exported to France with a price of $400 per ton.

Currently, grape dried seeds, oil seeds as well have high demand nationally and internationally and we are also having factories producing food seed oil.

Getting antimicrobial chemotherapeutic remedies, particularly in natural raw materials, is current and prospective permanent. Also of major importance is getting antifungal properties and antioxidative substances from plant raw materials. Research results on the activity of dissoluble etanon in water, allow production of preparations with pronounced biological activity, useful in medicine and veterinary use in combating bacterial diseases, fungal infections etc.

Currently, after processing the grapes are produced about 59 thousand tons of pomace. Marc contents of seeds make up 40–65%. So annually in the country, the grape processing plants can be harvested 20–25 thousand tons of seeds. Given the fact that enotatin content in grape seed makes up on average 8% annually could get about 2–2.5 thousand tons of enotanin. Prior research showed that leaching enotaninului yield is 65–70%. Thus, it may get about 1.2–1.3 thousand tons of active ingredient to changing them to produce medicinal preparations, veterinary, agricultural, etc.

Among all waste formed after the vinification process is the marc, which is the result of distilling alcohol from wine products. The perspective method of treating this waste concentrates worldwide is considered anaerobic digestion, which allows not only solve environmental challenges but also allows the conversion of organic pollutants into biogas, which can still be used as an alternative source of heat and electric.

The data presented above demonstrates that Moldova must build a complex factory for processing secondary wine products. An argument convincing enough to build such an enterprise serve the scientific data presented in the monograph [4].

2.1 Plastic Waste Management in the Republic of Moldova

The plastic materials represent about 10–12% of the amount of household generated waste in the country. Not all the categories of plastic materials are subject to recycling. The plastic materials in the majority of PVC degrade over several centuries, between 100 and 1000 years. Therefore we ascertain how great the pollution that they produce being persistent in time. The basic problems of recycling of plastics are in the variety of their frequencies, which are recycled separately and mixing them make it impossible to recover. The most often is considered optimal recovery separated from other plastic materials of the receptacles from PET by producers and consumers of soft drinks, which prevents mixing with other resins.

Beneficiaries dealing with plastics processing decide the categories of waste accepted for collection through specialized points.

For these reasons, it is necessary to organize the collection of differentiated plastics and training potential generators of waste plastic materials, including the general public. Concomitantly with the increasing amount of organic plastic materials in use, the problem of recycling is becoming more topical. Proceeding from the particularities of the composition of the waste polymers, plastics processing at low temperatures, which allowing us to obtain the polymer raw material (for example, construction materials, ramps, pillars, etc.) is the most promising.

Plastic deformation of solid bodies usually leads not only to changing its shape but also to the defects that change the physicochemical properties, including the reactivity. The accumulation of defects is used in chemistry to accelerate reactions involving the solid substances, for reducing the temperature processes and for other manipulations for increasing chemistry interaction in the solid-phase. The mechanical energy absorption initiates destruction of polymers, polymorphic transformations, and other reactions. In order to communicate the mechanical energy of plastic materials, the last are processed in extrusion reactors by using the crushers and conveyors. During of such a treatment, in addition to shredding polymers, are taking place the structural changes of substance, there are formed a lot of flaws and the material becomes reactive. In the simultaneous processing of several types of polymers, between them occur interactions–chemical reactions. However, analogous in the case of the thermal activation in the solid-phase reactions, for the initiation of mechanical-chemical reactions it is necessary to transmit the dust a sufficient quantity of mechanical energy. The energy can be communicated through the crusher and polymer screw reactor. The screws acting through the impulses on the polymer shredding. The strength and the result (the degree of transformation after the chemical reaction) of such actions depend on the acceleration of the movement of the screw within the reactor. Previously, the use of mechanical activation of substances in the production of new materials has been limited by the lack of reactors operating safely. Currently, such reactors exist and can be used in such a process harmless for the environment as a secondary plastic processing [7]. By a project of technology transfer, in the Republic of Moldova was implemented a complex process of waste plastics [8].

Waste of plastics separated from polyolefin (PE, PP), unwashed are introduced into the hammer mill where it is comminuted up to 50 mm parallel the polyolefin are crushing separately, and then they are crushing separately until the fraction 0.5 mm up to 10 mm (Fig. 4 and 5).

Comminuted and mixed waste in the proportion of 1/3 of the polyethylene and polypropylene: 3/4 of plastics waste with unidentified composition is loaded into the bunker with a speed of 0.2 m/s through the screw charger.

Fig. 4. Waste of plastics

Fig. 5. Waste of plastics comminuted by PP comminuted with containing unidentified

Thus, in the extruder (stainless reactor) the mixture is heated to a temperature of 180–2100 C. The speed of charging of comminuted waste of plastic materials are adjusted, for the temperature of self-heating of the mixture in the extruder (steel reactor) to vary in the range of 180–2100 C and no more.

At this temperature (polyethylene) (polypropylene) is liquidized and other compounds remain as viscous (polystyrene PS, polyvinyl chloride, PVC, polyethylene terephthalate, PET serve as compounds of liaison and polycarbonates, polyamides, etc. constitute the filler mixture. The obtained viscous mass is homogenized and transferred to the calibrator.

Calibrated mass of material is transported in the matrix pressing machine, where under the action of pressure from 3 bars is obtained the finite product: forms for plastic pallets (Fig. 6).

Respecting the same regularity of the technological process, but varying the report of homogenization of plastic waste comminuted in a proportion of 1/3 parts of polyethylene and polypropylene 3/4 of plastics waste with unidentified composition may be obtained new forms of the sidewalk (Fig. 6). To the

Fig. 6. New products obtained from waste of plastic mass

processing of 1,000 kg of the waste of plastics, may be obtained 950–970 kg of new products from plastics.

In the Republic of Moldova, in cooperation with the Academy of Sciences of Moldova for the first time, it was implemented a complex technology of processing from waste plastics with obtaining of new products. This technology creates no other waste and at the same time don't generate any wastewater (Fig. 7).

Fig. 7. The sector of complex processing of plastic waste

2.2 Infectious Medical Waste Management in Republic of Moldova

Medical wastes management activity is a part of nosocomial infection control in health facilities (IMS). Nosocomial infections are listed as special health issues (Regulation on the national epidemiological surveillance and control of communicable diseases and health events, Gov.Dir nr.951 of 25.11.2013).

According to the Statistical Yearbook of the health system for 2014 in Moldova (Right Bank) are 348 state medical institutions (legal entities) and 1029 are private.

There are 3 categories of infectious medical waste generators:

Large manufacturers: Republican, municipal, district and departmental hospitals; University of dispensary clinics, Medical Research Institute, Institute of

Forensic Medicine, Forensic Regional Services, Pre-Clinical Departments of the University of Medicine, Faculty of Medicine and Pharmacology, etc.

Medium manufacturers: Medical institutions of primary care clinics of hospitalization, blood transfusion centers, microbiological laboratories and clinics, nursing, medical and dental offices, private hospitals and clinics, etc.

Small producers: dental laboratories, rehabilitation centers, funeral services, vision care centers, acupuncture clinics, preventive medicine centers, private clinics, etc.

Among all categories of waste products of medical activity in medical institutions as a result of daily care activities, are the following categories of waste (Regulation of medical waste management no. 06.8.3.45 of 10.12.2001, approved by the Doctor Chief State Sanitary no. 05-00 of 14 December 2001 (Official Monitory of the Republic of Moldova, 2002, No. 13–15 art. No.:29).

"Hazardous waste" is the waste resulting from medical activities which present a real risk to human health and the environment, which are generated in the course of diagnosis, treatment, surveillance, disease prevention and rehabilitation, including medical research and production, testing, storage and distribution of medicines and biological products;

"Medical waste assimilated with domestic waste" are non-hazardous waste whose composition is similar to waste and no major risk to human health and the environment;

"Pathological waste and parts (parts) anatomical" these include tissues and organs, anatomical results of surgical instruments, the autopsy and other medical procedures; this category also includes animals used in laboratory diagnostic work, research and experimentation;

"Pathological waste and anatomical parts" are wastes include tissues and organs, anatomical results of surgical instruments, the autopsy and other medical procedures; this category also includes animals used in laboratory diagnostic work, research and experimentation;

"Infectious waste" are liquid and solid wastes containing or contaminated with blood or other body fluids and materials that contain or have come in contact with viruses, bacteria, parasites and/or toxins microorganisms;

"Chemicals and pharmaceutical waste" posed by chemicals are solid, liquid or gas that can be toxic, corrosive or flammable; expired drugs and chemotherapy residues of substances that can be cytotoxic, genotoxic, mutagenic, teratogenic or carcinogenic.

"stinging-cutting waste" is waste that can cause mechanical damage by puncturing or cutting; "Radioactive waste" is solid, liquid and gaseous waste resulting from nuclear activities medical diagnosis and treatment, containing radioactive material;

According to the World Health Organization, 80% of all waste resulted from medical activities are considered inoffensive and are managed as waste and other hazardous materials, 20% are infectious, toxic or radioactive.

The survey data conducted in 2004 with the support of the World Health Organization show that in Republic of Moldova the estimated amount of hazardous waste from medical activity is 0.05 kg/patient/day treated at home for hospitalized cases–average 0.44 kg/bed/day. Thus the total production of hazardous waste was estimated between 10 and 11 tons of waste per day 4000 tons per year.

Currently, the total amount of waste reported by medical institutions has increased, accounting for about 8.7 million tons per year, including infectious waste (syringes, infusion catheters, etc.).

In Chisinau daily it is accumulate about 3 tons per day of waste medical activity annually–about 1095 tons of wastes. At the same time, 120 million syringes annually and 6 million infusion systems are used throughout the country.

According to the World Health Organization, the estimated operational costs for treating 1 kg of infectious waste varies between 0.13–2.2 US $, depending on the applied method. The lowest costs for treating infectious waste is the method by autoclaving (0.13–0.36 US $/kg).

Academy of Sciences of Moldova and the Ministry of Health and Ministry of Environment in 2015 implemented a pilot project for treatment of infectious medical waste. It was developed an authorized treatment of infectious medical waste (Figs. 8 and 9).

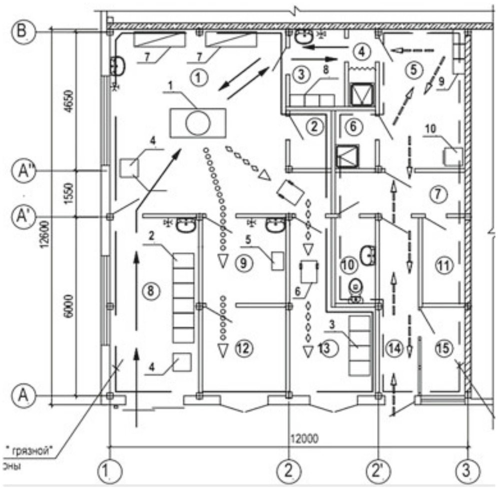


Fig. 8. Plan scheme of waste treatment point

In Fig. 8, 1 is working room of the plant; 2, 3, 4, 6, 7, 10, 11 are public spaces in the clean and dirty; 8 is room for receiving and temporary storage of waste (dirty area); 9 is room for washing and disinfecting bins; 12 is room for temporary storage of bins after washing and disinfection process; 13 is room for temporary storage of waste processed-disinfected; 5, 15 are rooms for personal and workplace.

Fig. 9. Authorized infectious medical waste processing point

STERISHRED technology and facility for treating hazardous waste from medical activities implemented at this enterprise, allows the obtaining of the final product environmentally friendly and health. Processing infectious waste takes place at low temperatures (136$\circ$C) with grinding, sterilization and destruction of confidential information storage units, in one compact space. The device converts waste "yellow bags" (biologically hazardous), including stinging and sharp objects in non-hazardous municipal waste-sterilized. Waste results, up to 70% can be processed at other companies for the obtaining of new products.

3 Conclusion

- Moldova has resources and scientific potential for capitalizing secondary wine products.
- Create a pilot plant for treating waste plastics complex served as an impetus for development of the field of waste recovery in Moldova.
- Academy of Sciences collaboration with the private sector has led to an authorized point of treating infectious medical waste, which resolved partially the pressing problem of neutralizing medical waste in Moldova.

References

1. Braunegg G, Bona R et al (2004) Solid waste management and plastic recycling in Austria and Europe. Polymer-plastics technology and engineering 43(6):1755–1767
2. Chisinau Statistica (2010) Anuarul statistic al republicii moldova
3. Damani N, Koolivand A et al (2013) Hospital waste generation and management in some provinces of Iran. Toxicological Environ Chem 95(6):962–969
4. Duca G (2011) Produse vinicole secundare
5. Ikeda Y (2016) Current home medical care waste collection status by nurse in Japan. J Air Waste Manag Association
6. Kumar S, Mukherjee S, Chakrabarti T, Devotta S (2007) Hazardous waste management system in india: an overview. Critical Rev Environ Sci Technol 38(1):43–71
7. Macaev F, Bujor S, Mereuţa A (2011) Reciclarea deşeurilor din mase plastice prin procedee mecanochimice
8. Macaev F, Mereuţa A, et al (2016) Procedeu de reciclare a deşeurilor de mase plastice/brevet de invenţie md 949 z

9. Sivapullaiah PV, Naveen BP, Sitharam TG (2016) Municipal solid waste landfills construction and management—a few concerns. Int J Waste Resources 6(214):
10. So WMW, Cheng NYI, et al (2015) Learning about the types of plastic wastes: effectiveness of inquiry learning strategies 44
11. Tugui T, Duca G, et al. (2013) Municipal solid waste composition study
12. Vyas P (2011) Municipal solid waste in India. J Ind Pollut Control 27(1):79–81

How to Popularize Green Residential Buildings in China: A Survey Study from Sichuan

Jun Gang[1(✉)], Dirong Xu[2], and Ying Wei[1]

[1] Sichuan Institute of Building Research, Chengdu 610003, People's Republic of China
467488122@qq.com

[2] College of Architecture and Urban-Rural Planning, Sichuan Agricultural University, Chengdu 610003, People's Republic of China

Abstract. Green buildings have been ushered in a period of rapid development in China. For there are few researches on economic benefit analysis, a result that most of Chinese people don't want to purchase residential buildings because of misunderstanding that green buildings have a higher cost. This paper focuses on how to popularize green residential building in China. In order to achieve the goal, the field survey is carried out, qualitative and quantitative descriptive methods are used to analyze the statement of green residential building in Sichuan Province. There are two important conclusions. Firstly, the incremental cost control of green building in our province is relatively reasonable. However, the incremental selling price is much higher than the incremental cost. Secondly, ordinary people can accept is much lower than the selling price. Finally, in order to improve the popularization of green residential building, some suggestions are provided.

Keywords: Green residential buildings · Survey study · Statistical analysis · Popularization

1 Introduction

China has been the world's largest carbon emitter and the world's largest energy consumer country since 2011. Building is one of the main contributor of energy consumption and pollution emission. China's building energy consumption has more than industrial and transportation, as the first source of energy consumption. To reduce the energy consumption, China had proposed the concept of green buildings in 2004 [7], which aim to save resources (including: energy, land, water and materials) and reduce the environmental impact. In the past five years, China's green building ushered in a period of rapid development. Green Building has first stated explicitly in the Five-Year (from 2011 to 2015) Plan of China [3]. In 2013, China's Green Building Action Plan has been launched [5]. And in 2015, the new version of "assessment standard for green building" was formally

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_108

implemented [2]. All these have brought positive results to promotion of green building. And also, more and more scholars have paid their attention to the green building in China, not only involving technology [4], but also containing cost [1] and management [8].

Although more and more peoples are showing solicitude for green building, but most of them are architects, engineers, scholars, and others related to the construction industry. Green buildings are not yet popularized since most common peoples don't still understand its concept and value. In China, all evaluation standards on green buildings don't contain economic indexes. The new version of "green building evaluation criteria" explicitly proposed that green building evaluation should take the economic analysis into consideration, but it did not put the economic analysis into the comprehensive evaluation system. Researches on economic benefit analysis are also seldom found [6]. As a result, most of Chinese people believe that green buildings (especially residential buildings) always have higher cost, so they don't want to purchase residential buildings. That an important reason why residential buildings cannot be popularized in China.

In order to understand the common peoples' attitude to green residential buildings fully, we did carry out an investigation to 34 green residential buildings and 935 peoples in Sichan in January, 2016. The following several sections will show the investigation data and some statistical analysis in detailed.

2 Survey Statement

This investigation mainly includes three aspects: (1) the public foundation of green residential buildings; and (2) the incremental cost and incremental price of green residential buildings; and (3) the people's acceptability to the incremental price of green residential buildings. Among them, the first part mainly surveys local residents' cognitive level and purchase intention about Sichuan's Green Buildings; the second part mainly investigates the incremental cost and the incremental price of green residential buildings in Sichuan; the third part mainly researches the public acceptability to the price of green residential buildings, and analyzes the people's needs and purchase intentions to green buildings in different hierarchy.

The questionnaire method was used to the ordinary peoples who are chosen randomly. This sampling survey was issued 1000 questionnaires, which 935 were recovered, and 901 are valid. The effective recovery rate of the questionnaires was 90.1%. The basic information of the respondents is as follows Table 1.

In addition, we also generally investigated all the 34 green residential buildings in Sichuan, which information is showed as Table 2.

3 Popularization Degree

3.1 Knowledge Level About Green Residential Buildings

In order to find out the popularized degree of green residential buildings in Sichuan, three questions are designed: (1) Do you understand the conception of

Table 1. Basic information of the respondents

Statistical variables	Sampling	Percentage (%)	Statistical variables	Sampling	Percentage (%)
Sex			Education		
Male	599	66.48	High school or below	104	11.55
Female	302	33.52	College graduate or bachelor	632	70.14
Profession			Bachelor or above	165	18.31
Self-employed	101	11.21	Monthly income RMB		
White-collar	119	13.43	7000	417	46.28
Ordinary worker	338	37.51	7000 14000	303	33.63
Student	273	30.3	14000 21000	99	10.99
Other	68	7.55	21000	82	9.1

Table 2. The basic information of the 34 green residential buildings

	One-star	Two-star	Three-star
Design identification	23	6	3
Operation identification	2	1	0
Total	25	6	3

green residential buildings? (2) Do you understand the functions of green residential buildings? (3) Do you know the policy of green building? The statistical results of the two questions are shown in Figs. 1, 2 and 3:

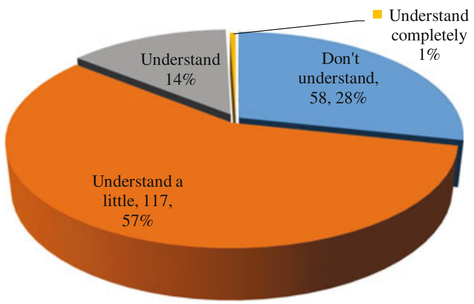


Fig. 1. People's understanding to the concept of green buildings

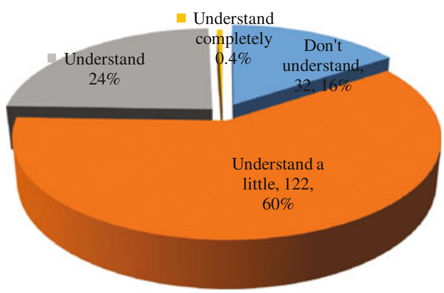


Fig. 2. People's understanding to the functions of green buildings

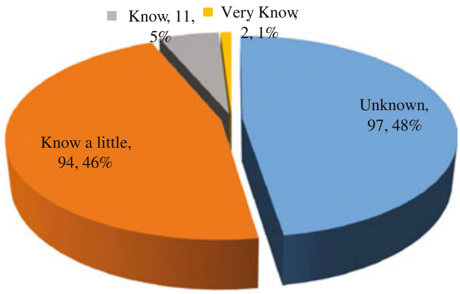


Fig. 3. People's understanding to the policies of green buildings

The survey results show that the people who know green buildings very clearly were accounted for only about 1%, some clearly only 14%; and understanding for the concept of green buildings, most people yet stayed in the "a little" and "a little more" dimensional level. Similar results are found about the functions and policies. 76% of the people don't understand or only understand a litter about the function of green residential buildings, while 94% of the people don't clearly understand the related policies about green buildings. However, in the process of investigation, we found that some people believe that the "green buildings" is the energy-saving buildings, and some people think "green buildings" is the buildings with many greenbelts. As a whole, the popularization of green residential buildings is serious shortage, although there is good policy support.

3.2 Purchase Intention to Green Residential Buildings

The purchaser of green residential buildings are the ordinary peoples, the government only plays a promotion role. Only when the public really understand what is the green building and how important to their live, it is possible to popularize green buildings. After analyzing the people's knowledge levels about green buildings, a question involving people's purchase intentions to the green buildings is designed, and survey results are showed in Figs. 3 and 4. It can be seen that 29.8% of the people think there is no necessary to develop green buildings in the first round of the investigation. However, after listening to the investigation team's explanation about the concept and functions of green buildings, 93.7% of people express their intentions to support the development of green buildings, which shows a good and extensive public foundation to popularize green residential buildings. Of course, it is necessary to enhance the publicity before popularization. Similarly, when investigated whether are willing to buy green residential buildings, after listening to the investigating group's carefully explanation, in the situation without considering incremental price, about 93.14% of respondents are willing to buy green residential buildings. In addition, nearly 7% of respondents are unwilling to buy green buildings, because of the worry about the incremental price. Indeed, the prices of green residential buildings will increase with the rise

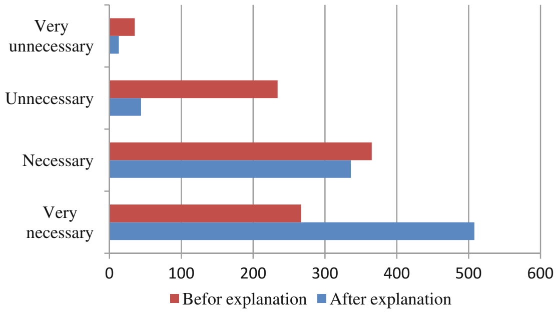


Fig. 4. Necessity of popularizing green residential buildings

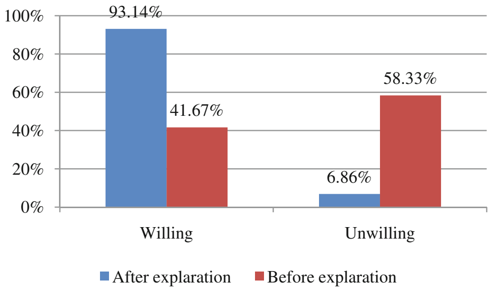


Fig. 5. People's willingness to buy green residential buildings

of design and construction cost. However, cost is not the unique factor which affects the selection to green buildings. It is found from existing literature that the main factors about the popularization of green residential buildings mainly include: government funding, comfort levels, crowding effects, cost factors and so on. Specifically, the survey found that most people give priority to benefit and comfort (36.82%). Followed by is the relevant government funding (29.1%). And 22.89% of the respondents tend to the benefit because of the diminishing cost in the operation phase. In addition, 10.2% of the respondents would buy green buildings if their relatives and friends have bought. In summary, the public has a high recognition for green residential buildings. They think it is necessary to vigorously promote and build the buildings, and also almost all are willing to buy this kind of building. This shows that the green buildings have a stable audience groups, there is a huge potential market to be tapped. But how to activate the vitality of the market, to stimulate people's actual purchase desire. The most important thing is that balance the relationship between the construction cost and the final benefit, that is, not only let people realize the bringing benefits of green residential buildings, and not simple to raise the price too much.

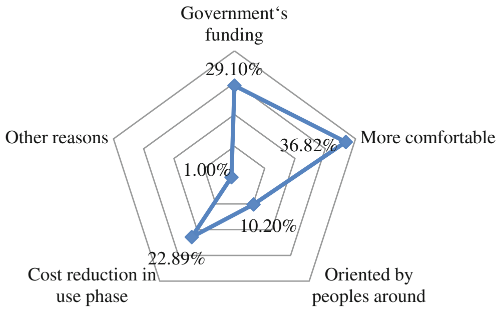


Fig. 6. Main factors affecting the purchase intention

4 Cost and Price

4.1 Incremental Cost and Incremental Price Statistics

From the statistics in Table 3, it can be seen that the incremental cost of green residential buildings in Sichuan is similar to that in China, which is controlled within 70 (CNY/m^2), the average incremental cost about 35 (CNY/m^2). It is less than 1% of total housing construction costs. It shows that the incremental cost of green buildings in Sichuan is reasonable.

Table 3. Incremental cost statistics

	One-star (CNY/m^2)		Two-star (CNY/m^2)		Three-star (CNY/m^2)	
	Average value	Scope	Average value	Scope	Average value	Scope
China	33	10 70	73	23 138	222	36 492
Sichuan	35	0–67	68	45–132	–	–

Notes: ① The nationwide data derives from the research results of "Post-assessment of green building in China", which is supported by the Ministry of Housing and Urban-Rural Development of China and led by of Chinese Society for Urban Studies; ② The data of Sichuan from the statistical results to 34 green residential projects from 2011 to 2015

At the same time, we also carried out statistics to the price of one-star green residential buildings in Sichuan, the results as shown in Table 4. It shows Sichuan Province's the average price of one-star green residential buildings is 7235 (CNY/m^2), and the average incremental price is 284 (CNY/m^2). Among them, the average price and average incremental price of the city of Chengdu are the highest. They are reached 9345 (CNY/m^2) and 385 (CNY/m^2), respectively. Compared the data in Tables 3 and 4, it can be found that the incremental price of green residential buildings in Sichuan is much higher than the incremental cost. The main reasons are as follows: (1) there is no a unified statistics standard to the incremental cost of green buildings in Sichuan, which leads the reported

Table 4. Incremental price statistics

	Actual selling price (CNY/m^2)		Incremental price (CNY/m^2)	
	Scope	Average value	Scope	Average value
Sichuan	4200–11500	7235	94–684	284
Chengdu	7800–11500	9345	136–684	385
Mianyang	4200–5100	4780	94–346	106
Nanchong	4900–5900	5360	115–526	218

incremental cost of some project seriously distorted; and (2) the current statistics to the incremental costs of green buildings only contain one-time investment cost, without considering the cost in the whole life cycle, such as cost in project planning, consulting, and operation stage. For example, although the consulting fees of green building must be produced for all the green buildings labeling programs, they are not counted in the reported incremental cost; and (3) for the same real estate, green buildings than other buildings often are occupied better natural resources (such as location, direction, greening, ventilation and etc.), and thus prices will be increased; and (4) housing prices than the cost of housing will usually retain a certain profit space, thus widening the direct gap between the price and the cost.

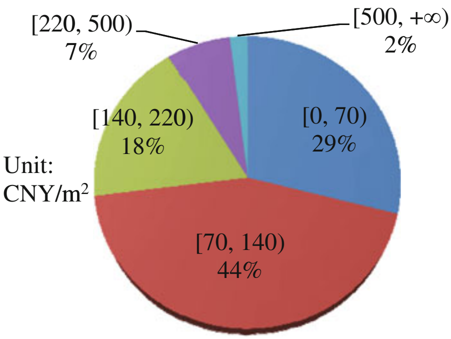


Fig. 7. People are willing to pay the proportion of the incremental price range

4.2 The Acceptability to the Incremental Price

In order to clarify the cost and efficiency of green residential buildings, a survey about public's acceptability to the incremental price is carried out. A question including five incremental price levels is set. The survey results are showed in Figs. 6–8. The results show that: (1) 73% of people are only willing to buy green residential buildings with increase price less than 140 (CNY/m^2); (2) only 8.8% of people can accept the increase price more than 220 (CNY/m^2). Thus, except

Mianyang, the incremental prices of green residential buildings are slightly high in Sichuan. Especially in Chengdu, the incremental prices of green buildings are much higher than the acceptable price of common people. Therefore, with enhancing the propaganda of green buildings, the price of green residential should also be reduced appropriately.

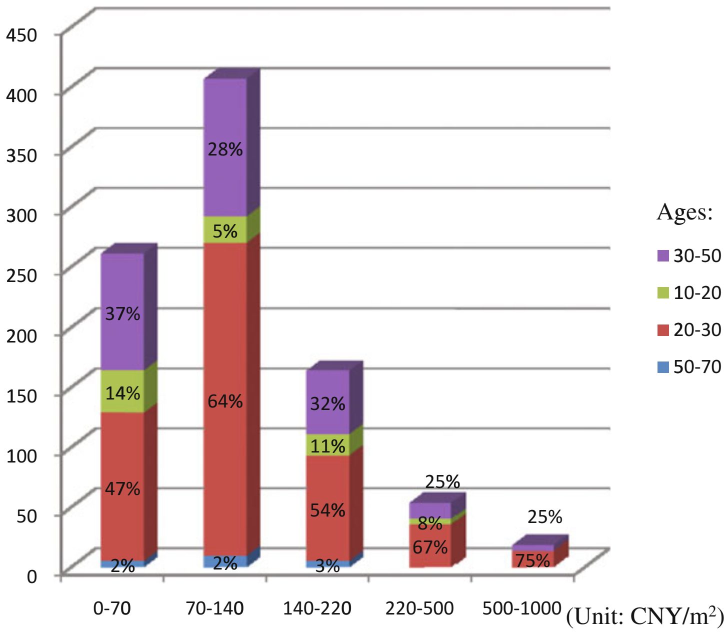


Fig. 8. Acceptability to incremental prices for different ages

5 Policy Proposals

In order to popularize green residential buildings, based on the survey results and statistic analysis above, the following policy recommendations are put forward.

5.1 Strengthen Publicity of Green Residential Buildings

From the above statistical analysis, it can be found that one of main reasons for hindering the development of green residential buildings in Sichuan is that the common people know too little about green buildings. The survey found that the vast majority of people are unclear about "what is a green building", "what can it bring what", "whether it increases the cost or not" and other core issues, which lead to a very low public acceptance for green buildings. At present, in China, only a small circle of people are working for green buildings. There are no specialized agencies or platform to popularize and propagate green buildings.

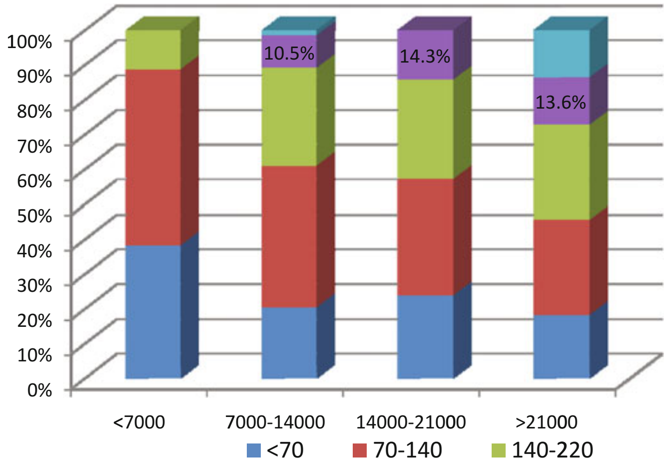


Fig. 9. Acceptability to incremental price for different income levels

Also, there is no corresponding funding support. Therefore, it is recommended that the government should increase budget to organize or guide some capable and responsible organizations or enterprises to give publicity to green buildings, including the conception, technology, production and benefit, to improve people's cognitive level for green buildings.

5.2 Build the Basic Database of Green Buildings

By the end of 2015, 3036 projects have obtained the identification of green buildings in China, and 56 in Sichuan. However, this information is only posted on few government websites, including the official website of ministry of housing and urban rural development of China, and some local official website. Moreover, these data are often published as public announcements, which are often too simple to understand. The project information in detail is always put in the attachment, and there is an open database to store them. As results, there are no channels to know and identify the green buildings' information, which hinder the further development of green building badly. So it is suggested that the government should establish official database on green buildings and open them to common people.

5.3 Drawing up Incremental Cost Statistical Standard

As describe above, the statistics on incremental cost of green buildings is variety in China. Unlike the German DGNB and Canada GBC, there is no strict requirements and uniform standard on the statistics of incremental cost in China. Some projects are counted rigorously, so reported relatively high incremental costs; however, some other projects are only counted minimal incremental cost items, some even no statistic, which lead to the unreasonable incremental cost. This brings serious difficult to correctly estimate the green buildings' costs and benefits, also to the management and service. In this situation, it is recommended that the unified standard about the incremental cost statistics of green buildings should be drawn up. In addition, the local government should require all green building projects must announce the data on incremental cost. Although this will increase cost slightly, it will pay an important role to popularize green buildings.

5.4 Enhance Cost and Price Control

From the above statistics analysis, it can be seen that the incremental price of green residential buildings in Sichuan was significantly higher than the affordable price of common people. Therefore, in order to promote the comprehensive development of green buildings, the price must be controlled. According to statistical data, the incremental price of green residential buildings should be controlled within 140 (CNY/m^2) as far as possible, especially for the one-star green residential buildings. In addition, the developers should also consider different demander according to the different price levels. For example, the projects which incremental prices are more than 140 (CNY/m^2) should focus on high-income groups. If the incremental price cannot be controlled to less than 500 (CNY/m^2), then the average monthly income of potential buyer should be more than 21000 CNY.

6 Conclusion

Although green residential buildings as the trend of the future development, many people still don't know exactly what's the green buildings. A green building not only means the vertical greening or roof garden, but also refers in the construction of the whole life cycle, to maximize the energy saving, land saving, water saving, material saving, environmental protection and pollution reduction. It is also called the sustainable development construction, ecological construction, energy saving and environmental protection construction etc. Looking around the property market in Sichuan, the green buildings are still very little. Three main reasons have been found according to the survey study in this paper. First, the common people know too little about green residential buildings. Second, the incremental prices of green residential buildings have exceeded the buyers' psychological expectations. Third, the environmental benefit and economic benefit for buyers are not widely accepted. Therefore, to popularize the green residential buildings, at the same time of improving the comfort, it is also necessary to strengthen the science propaganda and control the incremental price.

Acknowledgement. This research was supported by the special funds for building energy efficiency of Sichuan, and system science & enterprise development research center of Sichuan (No. Xq15C14).

References

1. Dwaikat LN, Ali KN (2015) Green buildings cost premium: a review of empirical evidence. Energy Buildings, 110:396–403
2. GB, T (2014) Assessment standard for green building. China Building Industry Press, Beijing
3. (2011) The twelfth five-year plan for national economic and social development of the People's Republic of China. Chin Nurs Res 25(3):207–214
4. Olubunmi OA, Xia PB, Skitmore M (2016) Green building incentives: a review. Renew Sustainable Energy Rev 59:1611–1621
5. The Ministry of Housing and Urban-Rural Development of the People's Republic of China (2013) Green building development plan. http://www.gov.cn/zwgk/2013-01/06/content_305793.htm
6. Yang Y (2015) Study on economic cost benefit evaluation of green building technology. Master's thesis, Southwest Jiao Tong University, Chengdu (in Chinese)
7. Ye L, Cheng Z et al (2013) Overview on green building label in China. Renew Energy 53(9):220–229
8. Ye L, Cheng Z et al (2015) Developments of green building standards in China. Renew Energy 73:115–122 (in Chinese)

Prognostics of Lithium-Ion Batteries Under Uncertainty Using Multiple Capacity Degradation Information

Fan Li[1(✉)] and Yusheng Wang[2]

[1] School of Economics, Sichuan University, Chengdu 610064, People's Republic of China
lifanscu@sina.com

[2] Uncertain Decision-Making Laboratory, Sichuan University, Chengdu 610064, People's Republic of China

Abstract. Prognostics of lithium-ion batteries play an important role in the intelligent battery management systems. The State of health (SOH) estimation for batteries often needs to be implemented under uncertain situations for the complicated operating conditions. In this work, a novel integrated approach based on a mixture of Gaussian process (MGP) model is presented for lithium-ion battery SOH estimation under uncertain conditions, where the degradation process parameters distribution can be learnt from the multiple available capacity measurements. To capture the time-varying degradation behavior, the distribution information of the degradation model parameter is extracted by fusing the training data from different battery conditions with the MGP model. Moreover, by exploiting the degradation model parameter distribution information, the PF algorithm is employed to predict the battery SOH. The correlation case study and comparison analysis are provided to show the efficiency and effectiveness of the proposed prediction method.

Keywords: Prognostics · Lithium-ion batteries · State estimation · Uncertainty

1 Introduction

With the advantages of high energy density, high galvanic, low-temperature performance, and long lifetime, lithium-ion batteries play a more important role in electronics system energy supply, and therefore have been widely used in communication, aerospace avionics, portable devices and other industrial areas [5,8]. However, battery deterioration and battery failure commonly occur, which can lead to a reduction in systems performance, result in increased costs and catastrophic failure [14]. Therefore, for the electronics device, prognostics and health management (PHM) has received increased attention in battery management system to determine the advent of systems failure and to mitigate system

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_109

risk through the evaluation of system reliability in terms of the current life-cycle conditions [13,16].

In order to evaluate the current health situation, some indicators such as the state-of-charge (SOC), the state-of-health (SOH) and the state of life (SOL) are commonly used to quantify the degradation of lithium-ion batteries [17]. To provide useful prognostics information in the reliability monitoring for battery systems health management, current researches have focused on SOC and SOH estimation for batteries. The main measurements required to effectively measure battery SOH are impedance and capacity [7]. In this study, battery capacity percentages relative to the initial capacity are adopted to measure the SOH.

In recent years, there are many valuable research works on the battery degradation modeling and prognostics for battery SOH estimation [9,15]. For example, neural networks and artificial intelligent methods have been rapidly used to estimate the battery SOH and predict the remaining useful life [11]. Stochastic filtering approaches such as Kalman filtering [1], extended Kalman filtering [4,10], unscented filtering and Bayesian filtering [2] are widely adopted in the SOH estimation of battery, in which the empirical degradation models are often used to build the dynamical system equation. Usually, the stochastic filtering methods based on the parameter process or degradation model can show good performance only when the degradation model used represents the actual system's behavior effectively. However, in practical battery conditions such as operating conditions, environmental conditions, and other complicated inherent system, it is difficult to obtain accurate state process model or parameter description when the uncertainties are considered. To address this issue, many approaches have been proposed such as a prognostic algorithm based on a Bayesian Monte Carlo method and the Dempster-Shafer theory for battery SOH estimation and RUL prediction [3]. Liu et al. [7] used the Gaussian process regression (GPR) to perform SOH prediction for lithium-ion battery, where the degradation trends are learned from battery data sets with the combination of Gaussian process functions. It should be noticed that uncertainty representation of the degradation model has not yet been fully investigated.

As mentioned above, it is urgent to develop the battery prognostics under the uncertainties. In this work, a novel approach to lithium-ion battery SOH estimation is presented through the integration of the MGP model learning and particle filtering. The proposed method consists of two phases, and the first is the MGP is used to learn the statistical properties of the degradation model parameter combining training data sets from uncertain battery conditions. Secondly, based on the parameter distribution information for the degradation process, particle filtering is exploited to obtain the battery SOH estimation. In the training phases, the GPR is exploited to initialize the distribution parameters for each component. Then, the MGP learning and the PF updating are recursively implemented. Finally, a case example is provided based on the NASA battery data sets to show the performance of the new prognostics method. The contributions of this study can be summarized in two points: the first is a fusion prognostics framework for the lithium-ion battery SOH estimation is developed by combining the degradation conditions from different batteries, the second is the proposed algorithm implements distribution learning for the multimode process under uncertainty.

The rest sections of this paper are organized in the following order: the GPR and the MGP model are introduced in Sect. 2. Section 3 details the proposed SOH prediction method which incorporates MGP learning and PF. Experiments and analysis are given in Sect. 4 to show the performance of proposed prognostics algorithm. Finally, we make the conclusions in Sect. 5.

2 Related Work

2.1 Gaussian Process Regression

The degradation model parameter vectors need to be treated as a dynamic process which captures the time-varying situations in the battery degradation cycles. To represent the system degradation behavior, The Gaussian process is considered. A stochastic process $\{g(x) : x \in \mathcal{X}\}$, indexed by elements from some set $\mathcal{X}$, is a Gaussian process with mean function $m(x)$ and covariance function $k(x, x')$. It means for any finite set of elements $x_1, \cdots, x_m \in \mathcal{X}$, the associated finite set of random variables $g(x_1), \cdots, g(x_m)$ have a multivariate Gaussian distribution, i.e.,

$$\begin{bmatrix} g(x_1) \\ \vdots \\ g(x_m) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_m) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_m) \\ \vdots & \ddots & \vdots \\ k(x_m, x_1) & \cdots & k(x_m, x_m) \end{bmatrix} \right).$$

It can be denote by $g(x) \sim GP(m(x), k(x, x'))$. Actually, the mean function and covariance function are defined as

$$\begin{aligned} m(x) &= E[g(x)], \\ k(x, x') &= E[(g(x) - m(x))(g(x') - m(x'))], \end{aligned}$$

for any $x, x' \in \mathcal{X}$.

Gaussian process represent distributions over functions, it provide a method for modeling probability distribution under multiple corruptions in complicated or uncertainty situations. When the accurately describe for the dynamical parameter process is hard to be obtained in advance, the Gaussian process regression (GPR) can be exploited to supply the approximation distribution of the parameter process through learning from the training data available [6].

Consider a set of training data $S = \{x_i, y_i\}_{i=1}^{N}$, the relationship between input x_i and output y_i can be modeled by $y_i = g(x_i) + \epsilon_i$, where ϵ_i is zero mean, Gaussian white noise with variance σ_n^2. From the GPR, if the prior distribution over $g(x_i)$ be assumed as Gaussian process, the posterior distribution over outputs conditioned on sample set S and the test input x_* is also a Gaussian process, and the mean and variance can be given by

$$\bar{g_*} = E[g_* | x_*, S] = k_*^T K^{-1} \mathrm{y}, \tag{1}$$

$$Cov(g_*) = k(x_*, x_*) - k_*^T K^{-1} k_*, \tag{2}$$

where, K is the covariance matrix whose entries is determined by the kernel function i.e. $K_{ij} = k(x_i, x_j)$. In order to capture the uncertainty of state process and the influence of noise, the kernel function can be expressed as $k(x_i, x_j) = k_g(x_i, x_j) + k_n(x_i, x_j)$, and the squared exponential kernel and constant kernel function are chosen as the covariance function of Gaussian process:

$$k_g(x_i, x_j) = \sigma_g^2 \exp\left[-\frac{(x_i - x_j)^2}{2}\right], \tag{3}$$

$$k_n(x_i, x_j) = \sigma_n^2 \delta_{ij}. \tag{4}$$

The parameters $\Theta = [\sigma_h, \sigma_n]$ are so-called hyper-parameters of the Gaussian process, which are actually the parameters of the covariance function. The hyper-parameters can be determined by exploiting numerical optimization techniques which is to maximize the log-likelihood function given by

$$\log p(y|X, \Theta) = -\frac{1}{2} y^T (K + \sigma_n^2 I)^{-1} y - \frac{1}{2} \log(|K + \sigma_n^2 I|) - \frac{N}{2} \log(2\pi),$$

where, I is the unit matrix of N dimensions, and N is the number of training data. Therefore, the learning process with GPR need to select the mean function and covariance firstly, then the hyper-parameters can be determined with the training data. After that, the prediction of the distribution parameter can be yielded with the new inputs.

2.2 Mixture of Gaussian Process Models

Consider the uncertainties in the degradation of different batteries, the available training data usually comes from several sources which may have different initial capacities or were collected from different measurement conditions. Therefore, to make full use of the various monitoring data, the MGP model is utilized to model the prediction distribution of the degradation model parameters by learning the mean and covariance from the multiple input space. With the obtained SOH measurements, the training set S is supposed to be composed of known K-components, $S_1, \cdots, S_K$, where $S_k = \{x_i, y_i\}_{i=1}^{N_k}$. For each training subset S_k, considering that there is a latent degradation model with uncertainty that can be used to describe the degradation trend for different battery components, the initial Gaussian processes are first used to learn the distribution parameters of the degradation process. Then, the MGP model which is obtained as a weighted combination of finite Gaussian process models can be exploited to fuse the distribution information from different components. Therefore, denote $p(z(x) = k)$ as the probability that sample x corresponds to the k component, then, with the mixture weights $\pi_i, i = 1, \cdots, K$, the posterior probability of the samples associated with the k^{th} component is given by:

$$p(z(x) = k|x) = \frac{\pi_k N(x|\mu_k, \Sigma_k)}{\sum_{i=1}^{K} \pi_i N(x|\mu_i, \Sigma_i)}. \tag{5}$$

Let $p(y_*|S_k, x_*)$ be the predictive pdf for the output variable y_* on the condition that the new sample x_* is obtained from the k^{th} component. Thus, the predictive pdf from the MGP models can be given as:

$$p(y_*|S, x_*) = \sum_{k=1}^{K} p(z(x_*) = k|x_*)p(y_*|S_k, x_*) \tag{6}$$

where, $p(z(x_*) = k|x_*)$ is the posterior probability that x_* belongs to k^{th} component, and it can also be denoted as $p(k|x_*)$.

Based on the structure of the MGP models, the state process for the battery degradation parameters can be given through the distribution learning which combines multiple possible degradation processes. It can be seen from Eq. (6), the predictive distributions from the MGP models can be determined from the posterior probability of the component the test sample belongs to and the associated distribution parameters update from the Gaussian process components when the new test sample is obtained. Commonly, the Expectation-Maximization algorithm, which is based on the maximization of the log-likelihood, is used for MGP model parameter optimization. To represent the time-varying process of the degradation model parameters in the training phase, recursive learning for the Gaussian mixture model can be exploited. In this way, denote the distribution learnt from k^{th} GPR component at time t as $\{\pi_k^{(t)}, \mu_k^{(t)}, \Sigma_k^{(t)}\}$, when the test sample $x^{(t+1)}$ arrives at time $t+1$, the mean and covariance of the k^{th} component, and its mixture weights can be updated as follows [18].

Step 1. Compute the posterior probability of the component:

$$p(k|x^{(t+1)}) = \frac{\pi_k^{(t)} N(x^{(t+1)}|\mu_k^{(t)}, \Sigma_k^{(t)})}{\sum_{i=1}^{K} \pi_i^{(t)} N(x^{(t+1)}|\mu_i^{(t)}, \Sigma_i^{(t)})}. \tag{7}$$

Step 2. Update the mixture weights:

$$\pi_k^{(t+1)} = \pi_k^{(t)} + \gamma(p(k|x^{t+1}) - \pi_k^{(t)}). \tag{8}$$

Step 3. Update the distribution parameters:

$$\mu_k^{(t+1)} = \mu_k^{(t)} + \gamma \frac{p(k|x^{t+1})}{\pi_k^{(t)}}(x^{t+1} - \mu_k^{(t)}), \tag{9}$$

$$\Sigma_k^{(t+1)} = \Sigma_k^{(t)} + \gamma \frac{p(k|x^{t+1})}{\pi_k^{(t)}}((x^{t+1} - \mu_k^{(t)})(x^{t+1} - \mu_k^{(t)})^T - \Sigma_k^{(t)}), \tag{10}$$

where, γ can be viewed as the control parameter which can be set as $1/N$ where N is the number of acquired samples.

3 The Proposed Method Based on MGP

The proposed method of SOH prediction has two phases. For the uncertainties in the battery capacity degradation, the MGP is able to combine the corresponding degradation process information under different conditions, thus, the MGP model is used to learn the appropriate distribution to represent the transition of the degradation model parameters in the training phase. In each mixture component, the nonlinear and time-varying parameter processes are treated as Gaussian process with associated means and covariance, which means the GPR is exploited for initialization. In the prediction phase, with the training data under the current condition, the trained MGP model is used to supply the importance density needed to produce the weighted samples in the Bayesian filtering implementation.

In fact, there are two parts in the training phase. Firstly, in order to capture the statistical properties for the degradation model parameters under uncertain state transition, the GPR is utilized to learn the distribution information from the different training sets. In each component, the training data can be given by $D_k = \{(x_i, y_i)\}_{i=1}^{N_k}$, where $y_i = \Delta x_i = x_{i+1} - x_i$, thus, the parameter at the next time instant can be expressed by adding the corresponding output to the current state, as the output with each input x_i is set as the change of the parameters. Then, for each component, the mean $\bar{g}_{N_k}$ and covariance $\mathrm{cov}(g_{N_k})$ are determined through the GPR Eqs. (1)–(2), and denoted $\mu_k^{(0)} = \bar{g}_{N_k}$ and $\Sigma_k^{(0)} = \mathrm{cov}(g_{N_k})$, which means that the initial mean and covariance are given in the Gaussian process components. Secondly, when the initial distribution parameters of MPG are given, the degradation model parameters pdf can be represented by using the recursive algorithm for MPG with the current training data. For the battery capacity data available from the current battery condition, denote the current training data $D_T = \{\Delta x^{(t)}\}_{t=1}^{T}$ as the parameter transition, where $\Delta x^{(t)} = x^{(t+1)} - x^{(t)}$, and $\Theta_k^{(t)} = \{p(k|\Delta x^{(t)}), \pi_k^{(t)}, \mu_k^{(t)}, \Sigma_k^{(t)}\}$ are the current distribution parameters for each component, then, the pdf of MGP can be computed from the MGP parameters updating using Eqs. (7)–(10). After that, the importance sampling and the resampling procedure are implemented from the standard particle filter algorithm. With the particles and associated weights, the battery capacity prediction $\mathrm{SOH}_{l+p}^{[i]}$ at p step after current cycle l can be computed by exploiting the degradation parameter samples $\tilde{x}_l^{[i]}$ for the i^{th} trajectory. Meanwhile, the SOH prediction at cycle l can be estimated by combining the N_s samples of capacity measurements with their corresponding weights. The steps of our proposed method are summarized as follows.

Step 1. Initialization:
Given the trained data set $S_L = \{(x_i, z_i)\}_{i=1}^{L}$ and the trained parameter set in different components $D_k = \{(x_i, y_i)\}_{i=1}^{N_k}$, $(k = 1, \cdots, K)$.

Step 2. Gaussian process regression-based learning in each component: For each component $k = 1, \cdots, K$, $\{\mu_k^{(0)}, \Sigma_k^{(0)}\}$=GPR$(D_k)$.

Step 3. Update the MGP parameters with current training data D_T:

Step 3.1 For each component$k = 1, \cdots, K$ $\Theta_k^{(t+1)} = \mathrm{MGP}(\Theta_k^{(t)}, \Delta x^{(t+1)})$

Step 3.2 Compute the density of MGP using Eq. (6)

Step 4. Denote $\mathrm{MGP}(\Delta x_{l-1}^{[i]}; \Theta^{(T)})$ as the MGP density with the parameters $\Theta^{(T)}$, and $N(z_l; GP_\mu(x_l^{[i]}, S_L), GP_\Sigma(x_l^{[i]}, S_L))$ represents the Gaussian pdf with the mean $GP_\mu(x_l^{[i]}, S_L)$ and the covariance $GP_\Sigma(x_l^{[i]}, S_L)$, then the importance sampling and resampling can be implemented:

Step 4.1 FOR $i = 1, \cdots, N_s$

$$\text{sample } x_l^{[i]} \sim x_{l-1}^{[i]} + MGP(\Delta x_{l-1}^{[i]}; \Theta^{(T)}),$$
$$\omega_l^{[i]} \propto \omega_{l-1}^{[i]} N(z_l; GP_\mu(x_l^{[i]}, S_L), GP_\Sigma(x_l^{[i]}, S_L)).$$

END FOR

Step 4.2 Calculate the normalization weights $\hat{\omega}_l^{[i]}$ and effective sample size N_{eff}

Step 4.3 IF $N_{eff} < N_{th}$

$$[\{\tilde{x}_l^{[i]}, \tilde{\omega}_l^{[i]}\}_{i=1}^N] = \text{resampling}[\{x_l^{[i]}, \hat{\omega}_l^{[i]}\}_{i=1}^N].$$

Step 4.4 IF $l < L$, let $l = l + 1$, turn to Step 4.1;

Step 5. Estimate the capacity with weighted samples $\tilde{x}_l^{[i]} = (a_l^{[i]}, b_l^{[i]}, c_l^{[i]}, d_l^{[i]})$:

$$\text{SOH}_{l+p}^{[i]} = a_l^{[i]} \times \exp[b_l^{[i]} \cdot (l+p)] + c_l^{[i]} \times \exp[d_l^{[i]} \times (l+p)],$$
$$\text{SOH}_{l+p} = \sum_{i=1}^{N_s} \tilde{\omega}_l^{[i]} \text{SOH}_{l+p}^{[i]}.$$

4 Case Example

4.1 Battery Data Set

In this section, a case study is conducted to validate the proposed approach. The battery capacity data used in this paper are collected by the NASA Ames' Prognostics Center of Excellence (PCoE) [12]. The lithium-ion batteries were run through different operational profiles, such as charge, discharge and impedance at room temperature. The battery capacity at cycles was adopted to measure the SOH, and the end-of-life criterion in the repeated charge and discharge cycles for the accelerated aging of the batteries was a 30% fade in rated capacity. Four batteries No. ~5, No. ~6, No. ~7, No. ~18 are acquired through the discharge with 2 A constant current level at an ambient temperature of 24 °C until the battery voltage fell to 2.7V, 2.5V, 2.2V and 2.5V, respectively.

To represent the empirical capacity degradation model, the conditional three-parameter capacity degradation model as follow was considered, where the constant variable κ was set at 2.

$$\text{SOH} = \tau \cdot \exp(\iota \cdot l) + \rho \cdot l^\kappa. \tag{11}$$

Using the Matlab curve fitting toolbox, the real degradation and the curve fitting for battery No. ~6 with the degradation model (11) are shown in Fig. 1. The empirical degradation model parameters are $(1.0263, 0.0037, 0)$, which were estimated by curve fitting with the capacity samples given from cycle 1 to cycle 100. It can be seen from Fig. 1 that the real SOH at prediction cycles are higher than the capacity degradation estimation. The root mean squared error is 2.8919 and the adjusted R^2 is 0.947, which are the goodness-of-fit statistics of the empirical model.

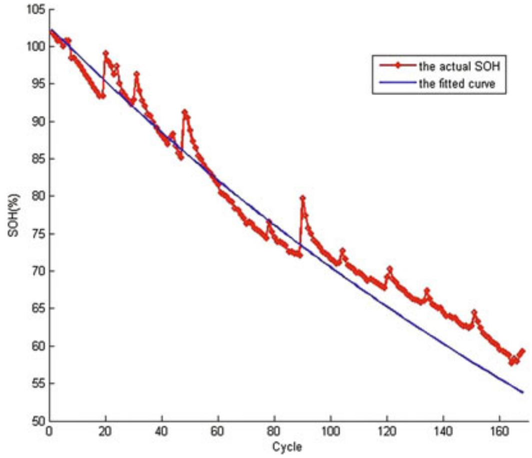


Fig. 1. The SOH of battery No. 6

4.2 The SOH Prediction

For the different operating profiles, the data from battery No. ~6 was assumed as the capacity samples under the current condition, and the data from other batteries, which represented the different operating conditions, were treated as the historical training data. In the training phase, there are two aspects. Firstly, with historical training data sets for different batteries, the GPR was exploited to learn the initial distribution parameters, where the mean function was chosen from the linear function $m(x) = ax + b$ and the covariance function was chosen as the exponential function expressed by Eqs. (3)–(4). Therefore, the hyper-parameters $\Theta = [a, b, \sigma_n, \sigma_g]$ were optimized with the maximization of the log-likelihood function. Secondly, the mixture density of MGP model is learnt by exploiting the current battery training set with the obtained distribution parameters from the different components. After that, the obtained distribution from data can be treated as the importance density, and then the particles and associated weights were updated using importance sampling and resampling procedure implemented as the standard particle filter algorithm.

In this study, the number of particles was set at 200. Consider the current battery No. ~6, the prediction began at cycle 100, which means the training data obtained from the cycle 1 to cycle 50 was used as the training set for the MGP learning, and the SOH measurements from cycle 51 to cycle 100 were used for the particles updating, and then the prediction results are shown in Fig. 2.

As shown in Fig. 2, the prediction combining three different training components can capture the degradation trends of SOH in most cycles. Therefore, from these results, the proposed method shows the effective prediction by fusing the multiple training data which represent the uncertain capacity degradation conditions.

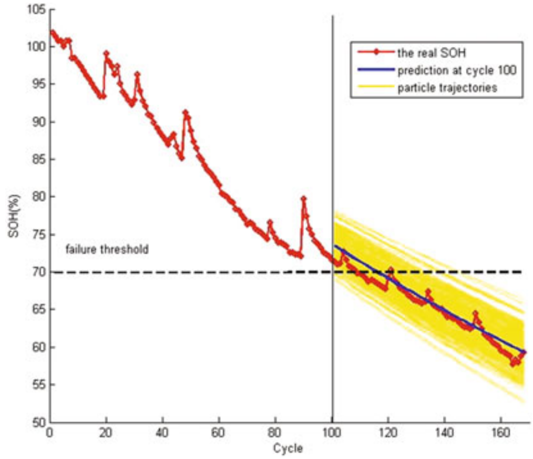


Fig. 2. Prediction results for battery No. 6 at cycle 100 with different training sets from batteries No. 5, No. 7 and No. 18

In addition, to further evaluate the proposed SOH prediction approach under uncertain condition, the method based on MGP learning and particle filtering (MGP–PF) was compared with the method which only considers the data from current battery. The GPR was used for parameter distribution learning, then the SOH estimation could also be implemented with the PF, thus the latter method denoted GPR–PF. Consider the current battery No. ~6, the prediction cycle was set at cycle 100, then the prediction results with two methods are shown in Fig. 3.

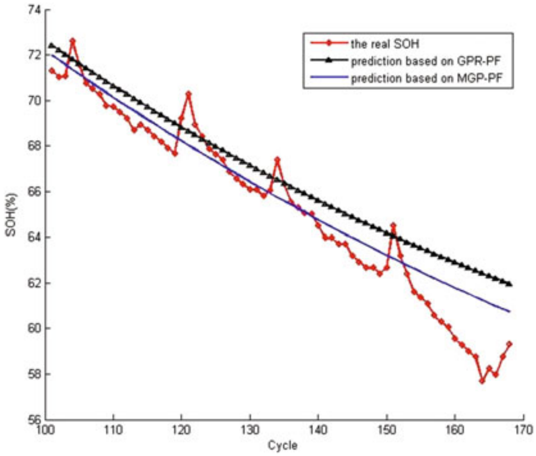


Fig. 3. Prediction results based on GPR-PF and MGP-PF for battery No. 6

For the battery No. 6, it can be seen from the Fig. 3 that the SOH prediction based on MGP and PF performed batter than the methods which only took the current capacity data into account. This indicates that the method which combines the priori information of capacity degradation process can improve the estimation performance in the current experiments. The reasons for this may lie in that distribution approximate which only considers the current measurements of battery ignores the useful fading information under different degradation conditions. Therefore, the method which fuses knowledge from multiple capacity degradation conditions is more suitable for lithium-ion battery SOH estimation with the uncertain battery degradation situations.

5 Conclusion

In this paper, a new model for the battery SOH estimation is presented based on the MGP model which was exploited to learn the distribution parameters from multiple training sets on the different degradation conditions. To represent the density of the degradation model parameters under uncertain conditions, the initialized distribution parameters can be updated recursively through MGP learning from the current capacity measurements of battery. Based on the degradation model parameter distribution information, importance sampling and resampling in the particle filtering framework was implemented. Our method is based on distribution learning from training data and does not assume any certain state model of degradation parameter, which is usually hard to be obtained in advance. Therefore, the proposed approach can also be applied to other estimation problems with unknown state space models under uncertain conditions.

Lithium-ion battery degradation under complicated conditions or working environments occurs in practice and different battery conditions can cause differences in the degradation. The prognostics for battery SOH estimation under uncertain conditions have many challenges, as different capacity degradation models with various degradation features cannot be obtained in advance. The uncertainties for the PHM of batteries is far more complex than the investigation in this paper, so further research on the implementation of fusion models under other uncertain conditions will be the focus of our future works.

References

1. Burgess WL (2009) Valve regulated lead acid battery float service life estimation using a kalman filter. J Power Sources 191(1):16–21
2. Cui H, Miao Q et al (2012) Application of unscented particle filter in remaining useful life prediction of lithium-ion batteries 2012:1–6
3. He W, Williard N et al (2011) Prognostics of lithium-ion batteries based on dempster-shafer theory and the bayesian monte carlo method. J Power Sources 196(23):10314–10321
4. Hu C, Youn BD, Chung J (2012) A multiscale framework with extended kalman filter for lithium-ion battery soc and capacity estimation. Appl Energy 92(4):694–704

5. Kim IS (2010) A technique for estimating the state of health of lithium batteries through a dual-sliding-mode observer. IEEE Trans Power Electron 25(4):1013–1022
6. Ko J, Fox D (2009) Gp-bayesfilters: Bayesian filtering using gaussian process prediction and observation models. Autonomous Robots 27(1):75–90
7. Liu D, Pang J et al (2013) Prognostics for state of health estimation of lithium-ion batteries based on combination gaussian process functional regression. Microelectron Reliab 53(6):832–839
8. Nishi Y (2001) Lithium ion secondary batteries; past 10 years and the future. J Power Sources 100(1–2):101–106
9. Piller S, Perrin M, Jossen A (2001) Methods for state-of-charge determination and their applications. J Power Sources 96(1):113–120
10. Plett GL (2004) Extended kalman filtering for battery management systems of lipb-based hev battery packs: Part 3. state and parameter estimation. J Power Sources 134(2):277–292
11. Qian K, Zhou C et al (2010) Temperature effect on electric vehicle battery cycle life in vehicle-to-grid applications. In: CICED 2010 Proceedings, vol 2010, pp 1–6
12. Saha B, Goebel K (2007) Battery data set. NASA AMES prognostics data repository
13. Vichare NM, Pecht MG (2006) Prognostics and health management of electronics. IEEE Trans Components Packaging Technol 29(1):222–229
14. Wakihara M (2001) Recent developments in lithium ion batteries. Mater Sci Eng R Reports 33(4):109–134
15. Williard N, He W, Pecht M (2012) Model based battery management system for condition based maintenance. In: Proceedings of the MFPT, vol 2012
16. Xu L, Xu J (2014) Integrated system health management-based progressive diagnosis for space avionics. IEEE Trans Aerosp Electron Syst 50(2):1390–1402
17. Zhang J, Lee J (2011) A review on prognostics and health monitoring of li-ion battery. J Power Sources 196(15):6007–6014
18. Zivkovic Z, Ferdinand VDH (2004) Recursive unsupervised learning of finite mixture models. IEEE Trans Pattern Anal Mach Intell 26(5):651–656

Analysis of the Effect of Low-Carbon Traffic Behavior on Public Bicycles

Jing Ma[1], Xin Liu[2], Zhineng Hu[1](✉), and Yong Cheng[2]

[1] Uncertainty Decision-Making Laboratory, Sichuan University, Cengdu 610065, People's Republic of China
huzn@scu.edu.cn

[2] Business School, Sichuan University, Chengdu 610064, People's Republic of China

Abstract. With increasing concern about $PM_{2.5}$, global warming, greenhouse gas emissions and the problem of the "last mile", the public bicycle system is gaining importance as viable mode in urban transportation worldwide. In order to more successfully promote the public bicycle to the residents, it is obviously necessary to understand in-depth what can affect the residents' usage of public bicycle. The study elicit the influence of demographic, attitude towards the behavior, subjective norm and perceived behavioral control on the usage of public bicycle. Based on the theory of planned behavior (TPB) and the binary logistic model, the paper conclude that the variable of gender, age, occupation and income are closely contacted with the usage of public bicycle. And the education qualification has a little effect on it. And there are significant relationship between usage public bicycle and environmental friendly, flexible, find-time, cleanness, the bicycle lanes, signals for cyclist, suitable weather.

Keywords: Public bicycle · Factor analysis · Low-carbon · Traffic behavior

1 Introduction

Though there are more and more convenient bus routes today in daily life, some back streets and new residential areas do not share public transport, and the residents often take more than 10 min to reach the bus station. To solve the public "last mile" problem, the community buses driving around the community entrance came into being in Chengdu in 2012. But because of the limitation of community buses, such as fixed lines and fewer bus stops, the residents were not satisfied with community buses, so the traffic pattern could not solve the problem perfectly. As the replacing for community buses, the public bicycles are more flexible and more popular.

Generally, the problem of traffic congestion and heavy emissions of CO_2 become more and more serious and people are increasingly worried about the $PM_{2.5}$ emissions and the smog problem in China. Therefore, low-carbon traffic

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_110

has been becoming increasingly popular in many cities. The local authorities have been paying more and more attention to public transport, and cycling in daily mobility is considered according to efficient means to reduce air pollution, traffic jams, and carbon emissions. Public bicycle systems (PBS) have turned out to be efficacious in promoting riding in many cities, particularly when contacted with public transportation.

Chengdu seems to be one of the most suitable cities for the development of shared bicycle in China due to flat terrain, mild weather, little sunshine and less wind as well as wide non-motor vehicle lanes. In addition, it is densely populated and there are about 238834 persons per square kilometers. Therefore, the public bicycle solve the Chengdu "last kilometers" of the problem effectively.

Despite public bicycle sharing system existing in many cities in China, cycling still does not become the habit of most residents, and researches for the influencing factors of PBS is not enough much and perfect. A study summed up the impact of spatio-temporal interactions on bicycle sharing system demand [10]. And a find pointed out bicycle sharing stations and greenness were motivators for bicycle commuting and public transport stations and elevation were deterrents for bicycle commuting by multinomial logistic regression models [6]. A research about public bicycle system in Beijing found that the bike-share choice is most sensitive comfort, user demographics, however, do not affect strongly on the bike-share choice, indicating the mode will draw users from across the social spectrum [4]. Firstly, this paper investigated the demographic factors related the usage of public bicycles. Then based on the theory of planned behavior the binary logistic model, this paper explored the factors that may affect the public bicycle selection of residents from the perspective of attitude towards the behavior, subjective norm and perceived behavioral control.

2 Literature Review

Attitudes and perception of an individual towards a particular mode would influence the behavior of that individual [8]. Commonly, the cycling requires a physical effort among the cyclist. The public bicycles are more and more popular today. The norm is also important for the usage of public bicycles. So it is necessary to understand which factors make a positive impact on the usage of public bicycles.

As the history of public bicycle system is not long, the researches on theory and practice of PBS are still immature. And most of the researches on public bicycle selection of residents concentrated on characters of public bicycle travel, the method or theory of travel choice and factors affecting public bicycle selection.

For the characters of public bicycle travel, a study found that the average travel distance of bicycle users is about 3.7 km, and the male users were 3.8 km, the female users are 3.6 Km through a survey towards the American travelers [1]. New findings unequivocally demonstrate that the mean journey length made by private bicycle is 700–800 m (0.44–0.5 miles) greater than those made by public bicycle [5]. And a survey found that the average distance traveled by bicycles for fitness was 14 km [9]. What's more, through the development of public bicycle planning in Hangzhou, Shi [16] pointed out that the public bicycle system facilities construction, the usage and operational mechanisms should be carried out to research from the perspective of public bicycle facilities. In addition, a paper that examines the tensions between the glamour of a global competitive city explained what roles might public bicycle paly and how might a public bicycle system business model operates if it were to serve the transport disadvantage [13].

For the method and theory of travel choice, based on the information theory, a selection procedure is proposed to select important variables in the logistic regression model, and the strong consistency of these programs is proved [2]. And Yang [17] built a structural equation model of resident's travel choice in the process of urbanization based on out chain. Sakari [12] put forward that public bicycle system should be viewed as part of public transport rather than as a separate cycling scheme. Furthermore, a study developed a methodology for categorizing bicycling environments defined by the bicyclist's perceived level of safety and comfort [14]. In addition, there is a research that develop an estimation method using sinusoidal model to fit the typical pattern of seasonal bicycle demand [11].

For the factors affecting public bicycle selection, it turned out that lower education and absence of docking stations within walking distance were associated with lower likelihood of awareness of the public bicycle system programs [3]. And a research explored the impact of specific weather conditions and calendar events on the spatio-temporal dynamics of the case study public bicycle sharing programs employing novel spatial analytical techniques [7]. Besides, there is the study agreed that residents' travel choices are mainly driven by environmental conditions and personal travel habits [4]. A survey outcomes in India revealed perceived benefits, physical barriers, safety hazards, social barriers, and road condition may be the major factor classes influencing bicycle mode choice used in exploratory factor analysis [15].

The current literature review reveals a dearth in studies that comprehensively elicit the influence of demographic, attitude towards the behavior, subjective norm and perceived behavioral control on the usage of public bicycle. Aptly, the current paper aims to identify the effect in those aspects. Moreover, the study also points out the determinants and correlates of public bicycle commuting.

3 Methodology

The factors were identified in the literature as influential in public bicycle system are numerous and diverse [15]. The study focuses on commuting trips in order to analysis the factors that influence people use the public bicycle based on theory of planned behavior (TPB). In this paper, the factors have been divided according to this rule that whether it is belong to "attitude towards the behavior", "subjective norm" and "perceived behavioral control", they are the three main variables that determine behavioral intentions. Where there is more positive attitude, the more support of important others, and the stronger the control of perceived behavior, there is more the behavior intentions about using public bicycle. And vice versa. The factors of personal, social and cultural (such as personality, intelligence, experience, age, gender, cultural background, etc.) indirectly affect attitude towards the behavior, subjective norms and perceived behavior control by influencing the behavioral beliefs, and ultimately affect behavioral intentions and behavior. Attitude towards the behavior, subjective norms, and perceived behavioral controls are conceptually distinct, but sometimes they may have a common basis of beliefs, so they are both independent and related.

We first investigated the factors of demographic by a simple distribution analysis specified for bicycle-shared and non-bicycle-shared commuters. Relevance analysis (e.g. Person's chi-square test) were applied to find the relationships between the factors of demographic and commuting by public bicycles or not.

After that, we investigated the determinants of public bicycle commuting with the model of binary logistic regression from the perspectives of attitude towards the behavior, subjective norms, and perceived behavioral controls.

4 Data Collection

4.1 Study Location

Chengdu is a large-size city of Sichuan province in the southwest of China with the resident population of 14.428 million. And it is about 604.1 square kilometers. Economic development is very rapid and it is one of the fastest growing cities in China. What's more, the environmental pollution is very serious, so it is important to develop low-carbon traffic in the city. The terrain of Chengdu is relatively flat so that the citizen develop riding habit is possible. And public bicycles became more and more popular in the city since 2016.

4.2 Questionnaire Survey

The questionnaire includes four sections. The first section focuses on the demographic variables, and the respondents were asked to report their gender, age, occupation, educational qualification and income. In the second, third and fourth section, the questions about attitude towards the behavior, subjective norms, and perceived behavioral controls are descripted in the questionnaire. Table 1 shows the characteristics about the questions.

The data for the study is collected usage face-to-face questionnaire survey by a sample collected using a random sampling method. And it was at subway station, university, attractions and business district we did this survey.

Table 1. Summary of questions

Attitude towards the behavior	Subjective norm	Perceived behavioral controls
(1) I think riding public bicycle is a cool choose to commute	(1) I do not use the public bicycle because I do not how it works	(1) I would not use the public bicycle because it is laborious
(2) I want to use the public bicycle because it is so flexible	(2) I would not use the public bicycle because the registration fee is too high(higher 300 yuan)	(2) I would not use the bicycle because there is not bicycle lane
(3) I want to use the public bicycle because it is time-saving	(3) I would not use the public bicycle because the unit price is too high(higher 2 yuan per hour)	(3) I would not use the bicycle because of heavy traffic
(4) I want to use the public bicycle because it is environmentally friendly	(4) I would not use the public bicycle if the destination is not near to my house(without 10 km)	(4) I would not use the bicycle because there is no signal for cyclist
(5) I want to use the public bicycle because it is entertaining	(5) I would not use the public bicycle if it takes too much time to find the bike.(beyond 5 min)	(5) I would not use the bicycle if it is night
(6) I want to use the public bicycle because it is good for health	(6) I would not use the public bicycle because the bicycle is not clean	(6) I would not use the bicycle if weather is not pleasant
(7) I would not use the public bicycle because it is not consistent with identity	(7) I would not use the public bicycle because the bicycle is not beautiful	
	(8) I would not use the public bicycle because I have private bicycle and e-bicycle	

4.3 Sample Characteristics

A total of 633 people are in the survey and 79% respondents have been used the public bicycle. Most of the respondents from various income classes use the public bicycle as transportation to school and to work. Before using the public bicycles, 73% of respondents used subway or bus as the main mode, 60% od them walked, 26% of them take a bike, 34% of them use car or taxi. In addition, it is at school and community sa well as subway and bus station where most respondents used the public bicycle.

5 Statistical Analysis

All data analysis were performed using SPSS. For the demographic data, we test whether it is independent between usage public bicycle and demographic character by Person's chi-square test (The significance level $\alpha = 0.05$).

Attitude towards the behavior that may affect propensity of using public bicycle were evaluated by using a binary logistic regression model (95% confidence). Participant characteristic were dichotomized before testing as potential confounders, including whether participants think it is cool, flexible, time-saving, environmentally friendly, entertaining, good for health, not consistent with identity. In addition, the potential confounders for subjective norm include the understanding about PBS, the registration fee, the unit price, the distance, the find-time, outward appearance, cleanliness, and whether the participants have private bicycle and e-bicycle. And the potential confounders for perceived behavioral controls include whether it is laborious, whether there is bicycle lanes, signals for cyclist, heavy traffic, visibility and weather. Final model fitness was checked with a (post-estimation) generalized Hosmer–Lemeshow goodness-of-fit test.

6 Results and Discussion

6.1 The Relationship Between Public Bicycle Selection and Demographic Factors by Chi-Square Test

When the demographic data was analyzed, Table 2 shows the sample distribution according to demographic factors. It reveals that the public bicycles are very popular among people. There are 503 respondents usage the public bicycle. The majority of cyclists who use public bicycle are women (65%). And for those who do not use public bicycle, it is equivalent between men and women. The Asymp.sig. (2-sided) is equal to 0.002, so it explain that the use of public bicycles is related to gender at the 99.5% confidence level.

For groups of aged 10–48, there is no significantly difference in column proportions between use and non-use. However, the proportion of the use aged 49–65 age (4%) is significantly lower than the non-use. What's more, among the use, the groups of aged 10–35 are significantly higher than the other age groups. The group mainly comprises students and the young worker who are more likely to take bikes to school or to work. In addition, the Asymp.sig. (2-sided) is equal to 0.006, so it explain that the use of public bicycles is related to age at the 99.5% confidence level.

For the group of students, it is obviously that the proportion of the use (78%) is higher than the non-use (49%). After all, the students are easier to accept new things, like public bicycle. And it is usually not too far from their home (dormitory) to school (classroom), so the bike is the best transportation for them. Besides, among the use, the groups of retirees (1%) are significantly lower than the other groups. The group mainly comprises elder who may feel more difficult to ride a bike. And the Asymp.sig. (2-sided) is equal to 0.003, so it explain that the use of public bicycles is related to occupation at the 99.5% confidence level.

For those whose income is less 3000, the proportion of the use is slightly higher than the non-use. But for those whose income is more than 3000, the proportion of the use is higher than the non-use. And for those who use the public bicycle, the groups whose income is less 1500 and between 1500 and 3000 are significantly higher than the other income groups. The group mainly prefer riding than driving in order to reduce the cost. What's more, the Asymp.sig. (2-sided) is equal to 0.005, so it explain that the use of public bicycles is related to income at the 99.5% confidence level.

For those whose educational qualification is at the same level, the proportion of the use and the non-use is equal on the whole. And the Asymp.sig. (2-sided) is equal to 0.345, so there is slight connection between educational qualification and the use of public bicycles at the 99.5% confidence level.

Table 2. Distribution of the sample according to demographic factors.

Variables	Total	Use	Non-use	Asymp.sig. (2-sided)
Gender	633(100%)	503(100%)	130(100%)	0.002
Male	241(61.9%)	176(35%)	65(50%)	
Female	392(38.1%)	327(65%)	65(50%)	
Age	633 (100%)			0.006
10–22	270 (42.7%)	227(45%)	43(33%)	
23–35	257(40.6%)	213(42%)	44 (34%)	
36–48	67(10.6%)	49(9%)	16 (5%)	
49–65	39(6.1%)	14(4%)	25(18%)	
Occupation	633(100%)			0.003
Student	444(70.1%)	391(78%)	63(49%)	
Office worker	138(21.8%)	93(18%)	45(36%)	
Self ?C employed	24(3.8%)	13 (3%)	11(8%)	
Retirees	13(2.1%)	4(1%)	9(7%)	
Income	633(100%)			0.005
≤1500	256(40.5%)	201(40%)	54(41%)	
1501–3000	184(29.1%)	155(31%)	29(22%)	
3001–6000	155(18.2%)	91(18%)	24(19%)	
≥6001	78(12.2%)	56(11%)	22(18%)	
Educational qualification	633(100%)			0.345
High school student	8(40.5%)	5(1%)	3(2%)	
College students	30(29.1%)	22(4%)	8(6%)	
Undergraduate	411(18.2%)	324(51%)	87(67%)	
Graduate or doctor	184(12.2%)	152(30%)	32(25%)	

Note: the Chi-square statistic is significant at 0.05 level

6.2 The Analysis of the Influencing Factors of Public Bicycle Selection by Binary Logistic Regression

The development of binary logistic model:

$$p = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)} = \frac{1}{1 + \exp(-\beta_0 - \beta_1 x_1 - \cdots - \beta_k x_k)}.$$

In the model, p is the probability of selecting public bicycles, x_1, $x_2, \cdots$, x_k are explanatory variables, β_0, $\beta_1, \cdots$, β_k are the coefficients of the corresponding explanatory variables.

For the factors of attitude towards the behavior, a regression analysis were gone on according to whether the use of public bicycle as the dependent variable, whether participants think it is cool, flexible, time-saving, environmental friendly, entertaining, good for health, not consistent with identity as the independent variable. For the factors of subjective norms, according to whether the use of public bicycle as the dependent variable, the knowledge of PB, the registration fee, the unit price, distance, find-time, cleanness, appearance, and private bicycle and e-bicycle as the independent variable, a regression model get. For the factors of perceived behavioral controls, regressed according to whether the use of public bicycle as the dependent variable, laborious, the bicycle lane, the heavy traffic, signal for cyclist, visibility and weather as the independent variable. Then we develop binary selection model and select independent variable by “input” method to research what factors affect the choice of public bicycle.

Table 3 shows that the final chi-square statistics is 10.483 after iteration from Hosmer and Lemeshow Check and the critical value: $\text{CHINV}(0.05, 8) = 15.507$, so the chi-square statistics is lower than the critical value. From the perspective of significance, 0.233 is higher than 0.05. It indicates that the model can be a good fit and there is no significant difference.

Table 3. Hosmer and Lemeshow check (attitude towards the behavior)

Step	Chi-square	df	Sig
1	10.483	8	0.233

Table 4 elicits the factors of significant contribution for chose of public bicycles about attitude towards the behavior. Among the attitudinal characteristics, the characteristics of “environmental friendly” and “flexible” turned out to be the prominent variables with the Exp (B) value of 2.626 and 2.190 when compared with the others. From the perspective of significance, the significance of “environmental friendly” and “flexible” are lower than 0.05 and the significance of constant is 0.008 lower 0.005, so it explains that there are significant relationship between using public bicycle and environmental friendly as well as flexible. Therefore, a binary selection model can be shown below as following.

Table 4. Binary selection model regression results summary (attitude towards the behavior)

Independent variable	B	S.E	Wals	df	Sig	Exp (B)
Environmental friendly	0.816	0.237	11.881	1	0.001	2.262
Good for health	0.123	0.231	0.283	1	0.595	1.130
Entertaining	0.229	0.212	1.163	1	0.281	1.257
Cool	0.238	0.240	0.985	1	0.321	1.269
Flexible	0.737	0.215	11.733	1	0.001	2.090
Time-saving	0.002	0.217	0.000	1	0.993	1.002
Not matching identity	37.232	68459278.282	0.000	1	1.000	1.478E16
Constant	0.619	0.235	6.930	1	0.008	1.858

$$p = \frac{1}{1 + \exp(-0.619 - 0.816x_1 - 0.737x_2)}.$$

In the model, p is the probability of selecting public bicycles, x_1 is 1 when the participants use the public bicycle because it is environmentally friendly and is 0 when they do not. And x_2 is 1 when the participants use the public bicycle because it is flexible and is 0 when they do not.

Table 5 shows that the final chi-square statistics is 19.108 after iteration from Hosmer and Lemeshow Check and the critical value: $\mathrm{CHINV}(0.05, 8) = 15.507$, so the chi-square statistics is lower than the critical value. From the significant, 0.064 is higher than 0.05. It indicate that the model can be a good fit and there is no significant difference.

Table 5. Hosmer and Lemeshow Check (subjective norms)

Step	Chi-square	df	Sig
1	19.108	8	0.064

Table 6 shows the binary selection model regression results of subjective norms factors which can contribute to an increased public bicycle used. Obviously, the characteristics of “find-time” and “cleanness” are more important than other variables because the Exp (B) value (2.262 and 1.775) of them is higher than others. clearly, the significance of “find-time” and “cleanness” are lower than 0.05 and the significance of constant is 0.000 lower 0.005, so it explains that there are significant relationship between using public bicycle and find-time as well as cleanness. Therefore, a binary selection model can be shown below as following.

Table 6. Binary selection model regression results summary (subjective norms)

Independent variable	B	S.E	Wals	df	Sig	Exp (B)
The knowledge of PB	−1.146	0.218	27.554	1	0.172	0.318
The registration fee	0.252	0.230	1.203	1	0.273	1.287
The unit price	0.139	0.273	0.257	1	0.612	1.149
Distance	−0.27	0.213	0.016	1	0.899	0.973
Find-time	0.816	0.237	11.881	1	0.001	2.262
Cleanness	0.574	0.212	7.347	1	0.007	1.775
Appearance	0.307	0.246	1.564	1	0.211	1.359
private bicycle and e-bicycle	0.416	0.333	1.564	1	0.211	1.516
Constant	1.019	0.200	26.044	1	0.000	2.770

$$p = \frac{1}{1 + exp(-1.019 - 0.816x_1 - 0.574x_2)}.$$

In the model, p is the probability of selecting public bicycles, x_1 is 0 when the participants do not use the public bicycle if he (she) can't find public bicycles in 5 min and is 1 others. And x_2 is 0 when the participants do not use the public bicycle because it is not enough clean and is 1 when others.

Table 7 shows that the final chi-square statistics is 11.125 after iteration from Hosmer and Lemeshow Check and the critical value: $\text{CHINV}(0.05, 8) = 15.507$, so the chi-square statistics is lower than the critical value. From the perspective of significance, 0.195 is higher than 0.05. It indicates that the model can be a good fit and there is no significant difference.

Table 7. Hosmer and Lemeshow check (subjective norms)

Step	Chi-square	df	Sig
1	11.125	8	0.195

In summary, Table 8 highlights the determinant variable of perceived behavioral control that is closed linked with the use of public bicycles. As shown, whether there are the bicycle lanes, signals for cyclist or suitable weather contribute much for usage of the public bicycles according to the value (2.088, 1.861, 1.760) of Exp (B). And the significance of them (0.001, 0.006, 0.007) and constant (0.001) is lower than 0.005, so a binary selection model can get as following.

$$p = \frac{1}{1 + \exp(-0.609 - 0.736x_1 - 0.621x_2 - 0.566x_3)}.$$

Table 8. Binary selection model regression results summary (subjective norms)

Independent variable	B	S.E	Wals	df	Sig	Exp (B)
The bicycle lane	−0.736	0.225	10.700	1	0.001	2.088
The heavy traffic	−0.179	0.299	0.359	1	0.549	0.836
Laborious	0.533	0.336	2.522	1	0.112	1.704
Signal for cyclist	0.621	0.228	7.449	1	0.006	1.861
Weather	0.556	0.210	7.241	1	0.007	1.760
Visibility	−0.090	0.209	0.185	1	0.667	0.914
Constant	0.609	0.187	10.635	1	0.001	1.838

In the model, p is the probability of selecting public bicycles, x_1 is 0 when the participants use the public bicycle if there are not bicycle lanes and is 1 when others. x_2 is 0 when the participants use the public bicycle if there are not signals for cyclists and is 1 when others. x_3 is 0 when the participants use the public bicycle if the weather is not suitable to ride and is 1 when others.

7 Conclusion

This study investigated the demographic factors related the usage of public bicycles, and also explored the factors that affect the choice of using public bicycles from the perspective of attitude towards the behavior, subjective norm and perceived behavioral control according to TPB. The following are the conclusion obtained from the study.

For the demographic factors, the variable of gender, age, occupation and income are closely contacted with the usage of public bicycle. And the education qualification has a little effect to it.

As for attitude towards the behavior, people are more concerned about the convenience and low-carbon. Therefore, there are significant relationship between usage public bicycle and environmental friendly or flexible. From the perspective of subjective norm, people maybe care about the time-saving and comfort more than appearance of the public bicycles so that they are more particular about find-time or cleanness. What's more, for perceived behavioral control, people are more concerned about safety, and it is important for people to be competent to cycle in some particular environment, so people would have a more positive perception about cycling if there are the bicycle lanes, signals for cyclist or suitable weather.

Acknowledgements. This research was supported by the projects XQ14B04 and SKZX2015-SB87.

References

1. Aultman-Hall L, Hall F, Baetz BB (1997) Analysis of bicycle commuter routes using geographic information systems: Implications for bicycle planning. Transportation Res Record J Transportation Res Board 1578(1578):102–110
2. Bai Z, Krishnaiah P (1991) Variable seleetion in logistic regression. J China Univ Sci Technol 21(2):23–31
3. Bernatchez AC, Gauvin L et al. (2015) Knowing about a public bicycle share program in montreal, Canada: are diffusion of innovation and proximity enough for equitable awareness? J Transp Health 2(3):360–368
4. Campbell AA, Cherry CR et al (2016) Factors influencing the choice of shared bicycles and shared electric bikes in beijing. Transp Res Part C Emerg Technol 67:399–414
5. Castillo-Manzano JI, López-Valpuesta L, Sánchez-Braza A (2016) Going a long way? On your bike! Comparing the distances for which public bicycle sharing system and private bicycles are used. Appl Geogr 71:95–105
6. Cole-Hunter T, Donaire-Gonzalez D et al (2015) Objective correlates and determinants of bicycle commuting propensity in an urban environment. Transp Res Part D Transp Environ 40:132–143
7. Corcoran J, Li T et al (2014) Spatio-temporal patterns of a public bicycle sharing program: the effect of weather and calendar events. J Transp Geogr 41:292–305
8. Davies D, Gray S et al. (2001) A quantitative study of the attitudes of individuals to cycling. Trl Report
9. Dill JJ, Gliebe (2008) Understanding and measuring bicycling behavior: a focus on travel time and route choice. Bicycle Lanes 29
10. Faghih-Imani A, Eluru N (2016) Determining the role of bicycle sharing system infrastructure installation decision on usage: case study of montreal bixi system. Transp Res Part A Policy Practice 94:685–698
11. Fournier N, Christofa E Jr, MAK (2017) A sinusoidal model for seasonal bicycle demand estimation. Transp Res Part D Transp Environ 50:154–169
12. Jäppinen S, Toivonen T, Salonen M (2013) Modelling the potential effect of shared bicycles on public transport travel times in greater helsinki: an open data approach. Appl Geogr 43(43):13–24
13. Jennings G (2015) Finding our balance: considering the opportunities for public bicycle systems in cape town, South Africa. Res Transp Bus Manag 15:6–14
14. Joo S, Oh C et al (2015) Categorizing bicycling environments using gps-based public bicycle speed data. Transp Res Part C Emerg Technol 56:239–250
15. Majumdar BB, Mitra S (2015) Identification of factors influencing bicycling in small sized cities: a case study of Kharagpur, India. Case Stud Transp Policy 3(3):331–346
16. Shi X, Cui D, Wei W (2011) The investigation of the relationship between using and planning of public bicycle in Hangzhou. Urban Studies
17. Yang LY, Zhu XN (2012) Travel mode choice in the process of rapid urbanization. China Soft Science

An Empirical Study on the Impact of Niche Overlap of Tourism Enterprise on Tourist Satisfaction

Shaojiang Lin[1], Jiaying Chen[2(✉)], and Jing Tang[2]

[1] College of Tourism, Sichuan Agricultural University, Dujiangyan 611830, People's Republic of China
[2] Business School, Sichuan Agricultural University, Dujiangyan 611830, People's Republic of China
cjysc@qq.com

Abstract. Vicious competition of tourism enterprises results from product homogeneity deeply manifested as homogeneity of tourism enterprise resource utilization, thus becoming the key impeding current tourism development. This paper uses niche theory as reference, illuminates the conception and connotation of niche overlap of tourism enterprise, and proposes its three-dimensional measurement model by interviews and questionnaire surveys of tourism enterprises, and this thesis also tests the data using methods such as EFA analysis and CFA analysis. On this basis, the influence mechanism of niche overlap of tourism enterprise on tourist satisfaction degree is analyzed by using structural equation. Verification result shows overlapping degree of space-time resource niche of tourism enterprises, niche overlap of market resource, niche overlap of ecology resource are all significantly negatively related to tourism satisfaction negatively.

Keywords: Tourism enterprises · Niche · Niche overlap · Measurement model

1 Introduction

Tourism industry is still in the initial phase of development at present, and the overall level of tourism enterprises' utilization efficiency of resources is relatively low, which leads to serious phenomena of product homogeneity and vicious competition affecting tourist experience. For tourism industry is a complex business ecology system which integrates many industries with blurred boundaries, problem of tourism enterprise competition cannot be explained if the competition is confined to traditional competition theory within the industry, thus many scholars have begun to use related theory of natural ecology system to research tourism phenomena, such as niche theory [8].

Niche theory, as basic theory of ecology, was widely applied in the research of related fields such as diversity of species, interspecies relation, etc. Niche refers

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_111

to a set of resources used in a community by a population or an individual, and is a "way of living" of itself [9]. Since the conception of niche is blurred, the theoretical cycle offers its measurement standard, namely niche measurements. And niche overlap is one of the most important indexes. Niche degree is used to measure the niche degree similarity of two (or more than two) populations, representing common use degree of them on a kind of resource to evaluate their level of competition [4].

In the application of the niche theory on tourism, foreign scholars prefer to use niche theory to study the tourism market segmentation. Lew and Duval [10] explained the Long Tail concept and proposed that the concept of the post-tourist, for example, is a Long Tail phenomenon. Pitts [12]'s study was to determine if there is a lesbian and gay sports tourism industry. Vanessa [15] used the Rocinha case, a favela in Rio de Janeiro, as a model for other communities to show the favela as a new niche market for tourism activity in Brazil.

Different from the foreign scholars, domestic scholars mainly discuss the definition and connotation of tourism niche, the construction of niche evaluation system and put forward niche optimization strategy from the perspective of tourism industry, tourism destination, tourism city and so on. Wang [18] firstly summarized the correlated theory research on the niche, and proposed that the industry cluster niche is its occupied resource space in the regional ecosystem through its own conditions, and further more brings forward the concept and features of the TIC niche based on its characteristics. And thus he held that the TIC niche consists of tourism resources, the cluster management and the cluster environmental niche. Wang et al. [16], by introducing the theory of ecological niche and based on the analysis of application, constructed the index system of the evaluation of theme parks in Pearl River Delta. Xiang [19] used Delphi method to choose and establish the tourism niche evaluation index, then evaluate the tourism niche of Zhangjiajie and Tianmenshan national forest park by means of on-the-spot investigation and questionnaire. Wang and Yu [17] proposed that the measurement system of the niche of regional tourism cities consists of 4 dimensions of tourism resources, market, socioeconomy and environment as well as 32 variables. Chen [3] explored how Taiwan could build a niche in Asia's cruise tourism industry.

The above researches mainly concentrated in the tourism industry, tourism market, tourism and regional tourism by using niche theory or thought. However, at present, the academic circle has little research on tourism enterprises, the core population in tourism business ecosystem. In fact, in the circumstance of explosive growth of tourism, vicious competition of tourism enterprises and product homogeneity have become the key impeding tourism's development. It is urgently needed to research the resource competition situation of tourism enterprises and resource utilization effect from the perspective of ecological system so as to optimize competition relationship.

The innovation of this paper is to applicate the niche theory to tourism enterprises, to construct a measuring index system of niche overlap degree of tourism enterprises, and to analyze the influence of niche overlap degree on tourist satisfaction by using SEM. This article mainly divides into the following several parts: the first part mainly introduces the research background and research progress, the second part mainly puts forward the theoretical model and research hypothesis, the third part constructs tourism enterprises measuring index system of niche overlap degree and tourist satisfaction, the fourth part is data collection and analysis which is based on the questionnaires and interview survey, the last part is the research conclusions and prospects.

2 Theory Model and Research Hypothesis

Scholars have identified niche factors from different aspects. People such as Hannan [6], Baum [1] elaborated the identification of niche from the aspects of enterprises' market resource demand and ability of providing products. They think enterprise niche, as the intersection of enterprise resource demand and enterprise producing ability, relies on the position where the enterprise is located and what it does. Yan A and Qing-Li DA proposed that enterprises' niche is formed by the interaction between other enterprises and themselves actively in special circumstances [21]. Shan M indicated that enterprise niche can be divided in three dimensions: human resources, tangible resources and intangible resources [13].

The above scholars' identification of niche can be summarized in two dimensions: market resource and human resource. However, tourism enterprises are service organizations satisfying tourists' tourism experience and demand by providing products in certain time and space, and the elements of tourism time and space of tourism enterprises are virtually tourism enterprises' resource. On this basis, this thesis proposes the research hypothesis as follows:

Hypothesis 1: Niche overlap of tourism enterprises can consist of three dimensions: space-time resource, market resource and productive resource.

Furthermore, in business ecosystem, competition within populations can not only affect their internal development, but also impact other populations' health. Tourists, as one of core populations in scenic business ecosystem, their health conditions are also certainly affected by the relationship of tourism enterprises. Presentation of tourist population health condition in the whole scenic business ecosystem is actually tourists' experience condition, which can be measured with the index of tourist satisfaction degree by us. Accordingly, this thesis proposes research hypothesis as follows:

Hypothesis 2: There is significantly negative influence of tourism enterprise niche degree on tourist satisfaction degree.

According to the above analysis, this thesis proposes theory study model and study hypothesis, as shown in Fig. 1.

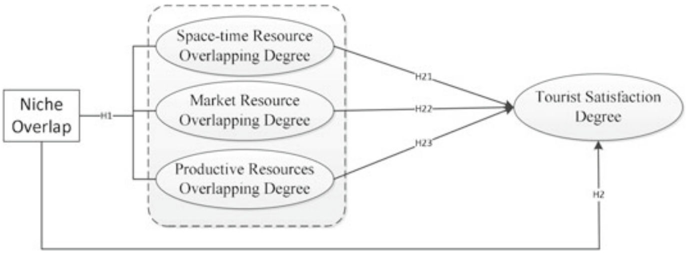


Fig. 1. Theory study model of "enterprise niche overlap-tourist satisfaction degree"

3 Variation Measurement

3.1 Measurement of Enterprise Niche Overlap

(1) Space-time Resource Overlapping Degree

Baum argued that enterprise niche, as the intersection of enterprise resource demand and enterprise producing ability, relies on the position where the enterprise is located and what it does [1]. Yan and Da [21] pointed out development opportunity and ecology strategy should be sought in time factor and social factor. Thus this thesis designs related measurement items to niche overlap, as shown in Table 1.

Table 1. Measurement items of space-time resource overlapping degree

Number	Measurement items	Sources
NO11	There are many enterprises of similar operation time-limit to your enterprise	Yan and Da [21]; Baum and Mezias [1];
NO12	There are many enterprises of similar life cycle phase to your enterprise	
NO13	There are many enterprises of same everyday business hours as your enterprise	
NO14	The enterprise is located in non-gold geographic area in the scenic spot	
NO15	There are many enterprises around your enterprise	

Data source: Arranged according to this research.

(2) Market resource overlapping degree

From the perspective of niche theory development, the influence of market resource dimension on enterprise niche has been emphasized more and more by scholars, and plenty of related research results have occurred in recent years. For example, scholars such as Dobrev [5], examined the effects of crowding in a market center on rates of change in organizational niche width and on organizational mortality. Hou et al. [7] considered that exorbitant overlapping between resources, abilities, markets, which populations need to exist and existing enterprises should be avoided to the greatest extent, for the niche overlapping will lead to competition of populations. Therefore, this thesis also adopts market resource dimension to measure the enterprise niche overlap (Table 2).

Table 2. Measurement items of market resource overlapping degree

Number	Measurement items	Resources
NO21	The enterprise is in bad situation in market competition	Hou et al. [7]; Dobrev and Hannan [5];
NO22	There are many enterprises of similar target customer groups to your enterprise	
NO23	Visitors think the product price of your enterprise is higher on equal basis of equality	
NO24	The sales method of the enterprise is similar to most enterprises	
NO25	The marketing channel of the enterprise is similar to most enterprises	

Data source: Prepared arranged according to this research.

(3) Overlapping degree of productive resources

Baum [1] considered that niche of enterprise relies on its location and functions, which is the intersection of enterprise resources and its productive capacity. Stuart and Podolny [14] proposed that niche of enterprises shall be identified by the number that the enterprise obtains in equivalent technical patents, while niche overlap in the enterprise shall be identified by the number of patents' relative references. According to the above analysis, the Thesis has extracted the dimensionality of productive resources to measure niche overlap in the enterprise as shown in Table 3.

Table 3. Measuring items of productive resources overlapping degree

No	Measuring items	Source
NO31	Staff's professional skills are better than other enterprises	Stuart and Podolny [14]; Hannan et al. [6];
NO32	Possessing similar techniques to most enterprises	Baum and Podolny [1]
NO33	Develop and launch new products and new services more often than other enterprises	
NO34	New products or services are quicker to be purchased by customers and seize the market	
NO35	The scale of manufactures is larger	

Data source: Sorted according to the research.

3.2 Tourist Satisfaction Degree

The research adopts a more mature SERVQUAL service quality scale to replace a simple tourist satisfaction evaluation. Proposed in 1990, SERVQUAL model included five dimensions, tangibles, reliability, responsiveness, assurance and empathy, and measured the distance of service quality by a questionnaire form [2]. At present, the model is considered as a most classic method to measure all kinds of service qualities appropriately and used in the relevant studies of all service industries [20]. In view of the particularity of different industries, the dimensionality and index of SERVQUAL model must be amended before it is applied to tourism industry [11]. Based on the previous studies, the paper has fine tuned the model and formed a tourist satisfaction evaluation system as shown in Table 4.

Table 4. Tourist satisfaction degree evaluation index system

Elements	Component items
Tangibles	1. Complete servicing facilities
	2. Modernized servicing facilities
	3. Staff in tidy clothes and appearances
	4. Comfortable and elegant decoration and environment
Reliability	5. Complete the things promised to the tourists without delay
	6. Show concerns and provide helps when the tourists encounter difficulties
	7. Provide products with reliable qualities
	8. Products safety can be trusted.
Responsiveness	9. Provide initiative and efficient products or service
	10. Provide prompt service
	11. Be always willing to help the tourists solve problems
	12. Provide service promptly and meet tourists' demands.
Assurance	13. Each staff is reliable
	14. Tourists will be at ease in service process
	15. The service attitude is passionate and polite
	16. Staff's communication skills
Empathy	17. Provide individual services for different customers
	18. Give an individual care for customers
	19. Pay attention to customer's demands
	20. Give priority to customer interests

Data source: Sorted according to the research.

4 Data Collection and Analysis

4.1 Data Collection

The research based on theoretical model and research hypothesis adopts one-to-one matching questionnaire method, chooses and designs each variable measurement scale and prepares and issues preliminary survey questionnaire. Due to indistinct tourism boundaries, it is difficult to define tourism enterprises. Thus, the selection of the research objects mainly focuses on the enterprises around scenic spots regarding tourism service as a main purpose. The research issues 450 questionnaires. Removing 38 questionnaires with incomplete data and obvious erring data, the research has recycled 402 valid questionnaires which indicate the recovery rate is 89.33%.

4.2 Testing of Validity and Reliability About Niche Overlap in the Enterprise

(1) Reliability test

The research adopts SPSS19.0 and AMOS17.0 software to test reliability and validity of scale. According to KMO value and Bartlett spherical test, the results show that KMO value is 0.917 and the significance probability of Bartlett spherical test's statistic is 0.000. The research adopts principal component analysis to leach factors, then extracts 3 factors in total with approximate 70.577% of gross number of variations according to the principle that characteristic root is greater than 1. The characteristic value of fourth factor is obviously less than 1, which indicates that the enterprise's niche overlap exists only 3 dimensionalities. Orthogonal rotation shall be used on factor reference axis and Varimax shall be used on the rotation. Loading coefficients of measuring items involved in all factors are basically greater than 0.7 after the rotation, which indicates that the scale has preferable discrimination validity and all measuring items are in line with the theorized expectations of factor structure. The above is shown in Table 5.

The Thesis adopts Cronbach's α coefficient to conduct a reliability test after the completion of factor analysis. Detailed analysis result is shown in Table 6.

It can be seen that internal consistency reliability of overall scale in official scale of the research is 0.917 and internal consistency reliabilities of all potential variable scales are all greater than 0.8 from reliability analysis. It indicates that consistency among all items of scales is comparatively high and official scale has a better internal consistency.

(2) Validity test

The research adopts SEM method to conduct construct validity test of all scales and the analytical model is shown in Fig. 2.

In 3 factors confirmatory factor analysis models of overlapping degree scale of niche, CMIN/DF value is less than 3, NFI, IFI and CFI values are all greater than 0.9, RMSEA value is less than 0.1 and fitting degree of all indexes is in better adaptation. The results have verified that niche overlap is formed by overlapping

Table 5. Rotating component matrix of niche overlap of tourism enterprise

Dimensionality	Measuring items	Factor loading		
		1	2	3
Space-time resources overlapping degree	NO11	0.198	0.186	0.798
	NO12	0.286	0.272	0.758
	NO13	0.284	0.217	0.795
	NO14	0.258	0.254	0.76
Productive resources overlapping degree	NO21	0.091	0.735	0.14
	NO22	0.132	0.765	0.103
	NO23	0.056	0.725	0.184
	NO24	0.165	0.729	0.27
	NO25	0.187	0.79	0.222
Market resources overlapping degree	NO31	0.84	0.133	0.23
	NO32	0.817	0.166	0.209
	NO33	0.851	0.151	0.2
	NO34	0.832	0.096	0.271
	NO35	0.867	0.141	0.194

Extraction method: Principal Component Analysis
Rotation method: Orthogonal rotation method with Kaiser standardization
The rotation is converged after 5 iterations.

Table 6. Official scale reliability

Scale (number of items)	Mean value	Variance of items	Internal consistency reliability
NO1 (4)	2.763	0.003	0.874
NO2 (5)	2.746	0.004	0.841
NO3 (5)	2.603	0.001	0.928
ST (4)	3.267	0.002	0.898
SRT (4)	3.18	0.002	0.867
SRS (4)	3.069	0.002	0.841
SA (4)	3.176	0	0.871
SE (4)	3.133	0.003	0.879
Total (34)	3.144	0.095	0.917

degree of space-time resource, overlapping degree of productive resource and overlapping degree of market resource.

Pearson's correlation coefficient is used in this study to calculate the relationship between niche overlap and tourist satisfaction. The empirical research indicates, there is no multi-collinearity between the variables, as shown in Table 7.

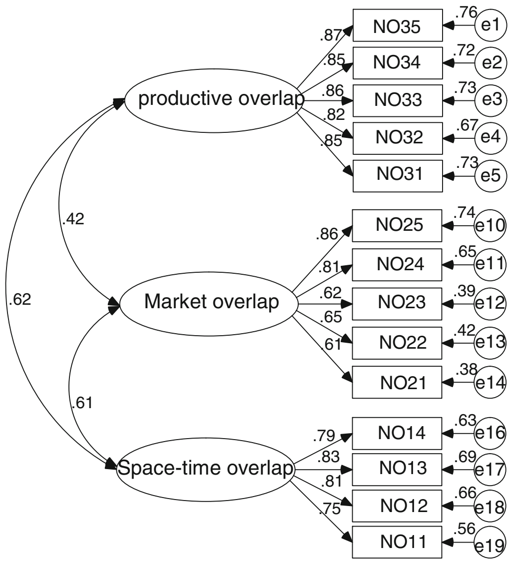


Fig. 2. Confirmatory factor analysis of niche overlap of tourism enterprise

Table 7. Pearson's correlation coefficient between niche overlap of tourism enterprise and tourist satisfaction degree

	Mean value	Standard deviation	NO1	NO2	NO3	TS
NO1	2.763	0.85794	1			
NO2	2.7462	0.70047	$.523^{**}$	1		
NO3	2.603	0.9468	$.558^{**}$	$.363^{**}$	1	
TS	3.1649	0.61748	$-.432^{**}$	$-.314^{**}$	$-.577^{**}$	1

Note: ** $p < 0{:}01$, NO1 - space-time resource overlapping degree, NO2 - productive resource overlapping degree, NO3 - market resource overlapping degree, TS - tourist satisfaction degree.

Based on results above, structural equation model is presented in this thesis to analyze the influence of overlapping degree of corporate niche on tourist satisfaction degree, and as overlapping degree of space-time resource, overlapping degree of market resource, overlapping degree of productive resource and tourist satisfaction degree are incorporated into the structural equation model, model fitting is well completed, as shown in Fig. 3.

As shown in analysis results above, overlapping degree of space-time resource, overlapping degree of market resource, overlapping degree of productive resource all have obvious negative influence on tourist satisfaction degree, with their standardized regression coefficient being $\beta = -0.11^{***}, \beta = -0.07^{***}, \beta = -0.3753^{**}$ ($p < 0.001$), respectively. It is thus clear that overlapping degree of corporate niche has negative influence on tourist satisfaction degree, among which the

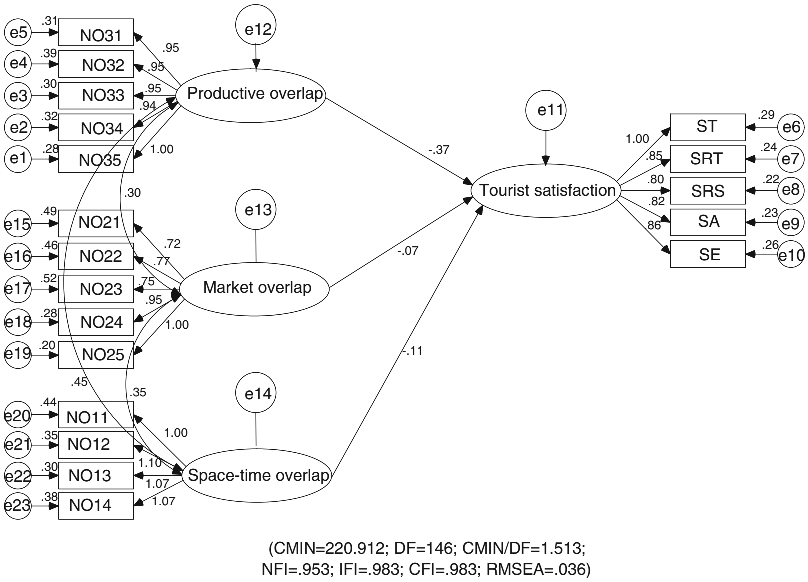


Fig. 3. The influence of niche overlap of tourism enterprise on tourist satisfaction degree

influence of overlapping degree of market resource is greater. Therefore, research Hypothesis *H2*, *H21*, *H22* and *H23* are all verified.

5 Research Conclusion and Prospect

The empirical research results indicate that overlapping degrees of corporate niche include overlapping degree of space-time resource, overlapping degree of productive resource and overlapping degree of market resource, and these three dimensionalities all have significant negative influence on tourist satisfaction degree. This is because: firstly, when business hours and location of corporations is the same as that of majority of corporations in business ecosystem of scenic spot, space-time resource overlap, as a result, competition in space-time resource of scenic spot between corporations is more intense, and in order to seek for more suitable time and space resource, the corporations will gradually consume their own resources in other aspects to survive in the competition.

Secondly, there are tourism enterprises of similar market niche inevitably in scenic business ecosystem, with different types of behavior occurring to fight for the same target customer group, such as price war, soliciting, etc., which will lead to decline of tourism satisfaction degree. Finally, tourists' experience of all products in the scenic area is actually experience of enterprise supply chains, so enterprises owning non-homogeneity technology are more likely to ally and integrate resources; and enterprises with homogenous technology tend to be subject to mutual repulsion, failing to provide better tourism experience for tourists.

In the future, research should be conducted further in combination with niche width of tourism enterprises, for niche overlap reflects the competition condition of resources used by tourism enterprises, while ecology width reflects the diversity degree of tourism enterprise resource utilization, and both of them interact with each other to influence tourists' tourism experience.

Acknowledgements. This work was funded by Youth Program of Social Science Research of Sichuan Province for the 13th Five-year Plan "Research on relationship embedded model between agricultural e-commerce enterprises and customer under the trend of industrial integration (SC16C022)", Project of Education Department of Sichuan Province "Research on the development choice between diversification or specialization of agricultural electricity supplier under the trend of industrial integration" and the National Natural Science Foundation of China (71501019).

References

1. Baum JA, Mezias SJ (1992) Localized competition and organizational failure in the manhattan hotel industry, 1898–1990. Adm Sci Q 37(4):580–604
2. Carman JM (1990) Consumer perceptions of service quality: an assessment of the servqual dimensions. J Retail 66:33–55
3. Chen CA (2016) How can taiwan create a niche in Asia's cruise tourism industry? Tour Manag 55:173–183
4. Colwell RK, Futuyma DJ (1971) Application of genetic algorithm to aircraft sequencing in terminal area. Ecology 52:567–576
5. Dobrev SD, Hannan MT (2001) Dynamics of niche width and resource partitioning. Am J Sociol 106:1299–1337
6. Hannan MT, Freeman J (1993) Organizational ecology. Harvard University Press
7. Hou J, Lu Q, Shi YJ (2011) Study on the variation and survival factors in the business evolution process based on organizational ecology. Ind Eng Eng Manag (IEEM) 5:1922–1926
8. Huang F (2001) Eco-principles' application in the optimum of tourist system. Ecol Econ 11:19–20 (in Chinese)
9. Hutchinson GE (1957) Concluding remarks: population studies animal ecology and demography. Cold Spring Harb Symp Quant Biol 22:415–427
10. Lew AA, Duval DT (2008) Long tail tourism: new geographies for marketing niche tourism products. J Travel Tour Mark 25:409–419
11. Lin S, Ren P et al (2014) Research on evaluation of travel service quality for jiuzhaigou valley on the basis of servqual. Wit Trans Inf Commun Technol 46:2157–2162
12. Pitts BG (1999) Sports tourism and niche markets: identification and analysis of the growing lesbian and gay sports tourism industry. J Vacat Mark 5:31–50
13. Shan M, Li G, Chen D (2006) Research on enterprises' competitive strategies based on ecology niche theory. Sci Sci Manag S & T 3:159–163 (in Chinese)
14. Stuart TE, Podolny JM (2007) Local search and the evolution of technological capabilities. Strateg Manag J 17:21–38
15. Vanessa OM (2014) Favela tourism: a new niche to be developed by Brazil. In: Eurochrie conference "hospitality and Tourism Futures", Dubai 6–9 October

16. Wang G, Dong GZ, Zhao JL (2008) Research on competitive pattern of theme parks based on ecological niche: take pearl river delta as an example. Tour Trib 12:45–51
17. Wang QR, Yu G (2008) A study on the measurement of the niche of regional tourism cities and cooperation-competition model-a case study on the Pearl River Delta region. Tour Trib 3:50–56 (in Chinese)
18. Wang ZF (2009) Research on the niche strategy of tourism industry cluster. Hum Geogr:12–15
19. Xiang YP (2009) A study on the evaluation of the tourism niche-case of Zhangjiajie and Tianmenshan national forest park. Issues For Econ 2:185–188 (in Chinese)
20. Xu M, Yu J (2001) Application of servqual scale to measure service quality. Ind Eng Manag 6:6–9 (in Chinese)
21. Yan A, Da Q (2005) Study of enterprise' niche and its motile selection. J Southeast Univ Philos Soc Sci Ed 7:62–67 (in Chinese)

Long Term Scheduling for Optimal Sizing of Renewable Energy Sources for Hospitals

Shabnam Mahmoudzadeh Vaziri[1], Babak Rezaee[1(✉)], and Masoud Amel Monirian[2]

[1] Faculty of Engineering, Industrial Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran
babak_rezaee@aut.ac.ir

[2] Masoud Amel Monirian, Nian Electronic Company, Mashhad, Iran

Abstract. Energy use is on the rise due to population increment and technology development. Generating electricity energy with fossil fuels causes environmental problems; moreover, their cost is high because of their depletion. For solving energy crisis and environmental problems, alternative methods for energy generation should be considered such as Renewable energy sources (RESs) which are considered as a promising alternative for fossil fuels. On the other hand, hospitals are one of the major consumers of energy. Energy costs consist a large amount of hospitals' costs; therefore, hospitals aim to minimize these costs owing to competition among health centers. The objective of this paper is to propose a long term scheduling model for the optimal sizing of the RESs among candidate technologies to meet the hospital's demand while minimizing the costs and maintaining reliability subject to economic and technical constraints. The model is run with real data from a hospital in Iran.

Keywords: Long term scheduling · Optimal sizing · Renewable energy sources · Hospitals

1 Introduction

Energy use has increased in last decades due to population and per capita usage increment. Due to fossil fuel sources depletion, importance of environmental issues and decrease in carbon emission, the energy crisis is on the rise. According to the energy crisis, most of the countries intend to generate energy with alternative sources and change their energy portfolio. Renewable energy sources (RESs) can reduce green gas emissions and environment pollution; therefore, they are proper alternatives for current generation method, and some researches focused on optimal sizing of RESs with different constraints. Moradi et al. [5] considered renewable energy sources as a reliable source of electricity. In this paper, they used Net Present Value (NPV) as an economic factor to determine type and capacity of sources. Sharafi and Tarek [8] presented a multi-objective model for determining the optimal size of renewable energy sources. They used ε-constraint

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_112

method to minimize total cost, fuel emission, and unmet load. Yamada and Raji [2] presented a model for determining the size of wind turbines, batteries, and solar panels. The goal of their paper was to minimize the costs in three years time horizon. Energy balance constraints, battery constraints, inverter model and load model constraints were considered in the paper. Abd-el-Motaleb and Bekdach [1] proposed stochastic model for optimizing size of wind turbines and energy storage units. They considered energy balance constraints, storage constraints, and also reliability constraints. Theo et al. [9] presented a linear optimization model for a hybrid power system in order to minimize net present value of costs. They chose wind turbines, solar panels and battery in an eco-industrial park as a case study.

Health centers and hospitals are one of the major consumers of electricity, where is necessary to generate energy in a continuous way and without blackouts. Health care forms a significant cost for governments; however, thanks to energy development, health care costs can be reduced [12]. Health care systems are evolving rapidly, and the number of patients is increasing due to emergence of new diseases and aging in society. Consequently, using modern medical equipment and patient increment cause to the rise in energy use. The number of health centers is increasing, and they evolved as a business industry because their competition increases. In competition market, patients have to pay parts of the costs, so health centers intend to decrease their costs as far as possible. Since hospitals consist 31% of health care costs, hospitals are considered in this paper. One way for reducing costs in hospitals is installation of RESs for energy generation [11]. Many countries allocate incentives to motivate investors for using RESs. Therefore, hospitals can use government's subsidy for supporting renewable energy sources. In this paper, hospitals are considered as energy consumers. Although supplying efficient energy is essential in hospitals, there are few researches on energy use in hospitals. Some scholars research about heating and cooling systems in hospitals [3,6]. Mavrotas et al. [4] developed a bi-objective model for minimizing costs and maximizing demand satisfaction. The technologies considered in their model were combined heat and power unit for providing power and heat, an absorption unit and/or a compression unit for providing a cooling load. Other papers focused on proper equipment selection in hospitals are [7,13].

The goal of this article is to determine optimal sizing of RESs in order to minimize the costs and conserve reliability. The paper is organized as follows. Section 2 proposes problem definition and mathematical model. Section 3 presents the numerical example, analyses results, and discusses sensitivity analysis. Finally, a conclusion is presented in Sect. 4.

2 Problem Definition

Since a significant amount of costs in a hospital is related to energy costs, careful energy planning is of a special importance in hospitals. In order to minimize the costs and reduce greenhouse gas emissions, hospitals can use renewable energy sources. Due to the importance of environmental issues, and consumer motivates for using RESs, governments consider some incentives for consumers producing renewable energies. Renewable energy sources are dependent on the natural situation for generation, and they are unpredictable; consequently, one way to increase their generation reliability is to use two or more renewable sources simultaneously [10]. For increasing generation reliability, wind turbines and solar panels are used in this paper, which means that hospital, according to its budget, can invest in solar panels and wind turbines for energy generation. Some part of required energy can be produced by RESs in this paper, and other parts are supplied by the grid. Moreover, the hospital can sell its extra produced energy to the grid. Economic factors such as inflation rate and interest rate affect decisions, so aforementioned factors are considered in this model by using NPV method for calculating objective function. The purpose of the model is to minimize investment costs in a specific period while maintaining RESs system's reliability. By solving the model, questions like the optimal size of wind turbines and solar panels and the optimum time for investment are answered.

Some assumptions are considered for the mathematical model as follows:

- The planning horizon is ten years which is divided into 120 months.
- According to geographic location, the hospital can invest in wind and solar energy.
- The hospital is connected to the main grid. Purchased energy price from the grid is varied in different time slots.
- Extra generated energy by RESs is sold to the grid with fixed price.
- Simultaneous buying energy from the grid and selling energy to the grid are not allowed.
- The investment cost of RESs is varied in time slots which is increased or decreased by the inflation rate.
- Interest rate and inflation rate are known and stable.
- The specific amount of money is considered as a budget for each time slot. The extra budget does not transfer to next time slot.
- Since the goal of the hospital is to focus on medical problems, the allocated area for installing RESs is limited.
- It is assumed that renewable energy sources use different rate of their capacity in each time slot $(\beta s, \beta w)$.
- In each time slot, minimum amount of energy has to be purchased from the grid to be used in emergency situations.
- For motivating hospitals to invest in renewable energy sources, government pays $a\%$ of sources investment cost.

To model the problem, the following notations are defined:
Indices:
t : Planning horizon (month) $t = 1, 2, \cdots, T$.

Parameters:

cap : Capacity of 1 kW of solar panels and wind turbines in a month;
βs_t : Utilization percentage of 1 kW of solar panels at time t;
βw_t : Utilization percentage of 1 kW of wind turbines at time t;
fs : Required area for installing 1 kW solar panel (Square meters);
fw : Required area for installing 1 kW wind turbine (Square meters);
F_s : Total area for solar panels (Square meters);
F_w : Total area for wind turbines (Square meters);
d_t : Energy demand of hospital at time t;
r_t : Reliability at time t;
$s_{\max}$: Maximum amount of allowable energy which can be bought from the grid;
$z_{\max}$: Maximum amount of allowable energy which can be sold to the grid;
p_s : Investment cost of 1 kW solar panel;
pm_s : Maintenance cost of 1 kW solar panel;
p_w : Investment cost of 1 kW wind turbine;
pm_w : Maintenance cost of 1 kW wind turbine;
c_t : Cost of buying energy from the grid at time (cent);
w : Price of selling energy to the grid (cent);
I_t : Budget at time t;
ss : Energy bought for emergency situations;
μ : Interest rate in a month;
α : Inflation rate in a month;
a : Government subsidy.

Decision Variables:

x_t : Installed capacity of solar panels at time t (kW);
y_t : Installed capacity of wind turbines at time t (kW);
z_t : Amount of energy bought from the grid at time t (kW);
s_t : Amount of energy sold to the grid at time t (kW);
u_t : Binary variable (1 if energy has sold to the grid).

Scheduling model intends to determine optimal sizing of RESs during 10 years time horizon as follows,

$$\min \sum_{t=1}^{T} \left[\frac{(p_s + pm_s)(1-a)(1+\alpha)^t x_t}{(1+\mu)^t} + \frac{(p_w + pm_w)(1-a)(1+\alpha)^t y_t}{(1+\mu)^t} + \frac{c_t(z_t + ss)}{(1+\mu)^t} - \frac{w s_t}{(1+\mu)^t} \right] \tag{1}$$

Subject to

$$\beta s_t \times \mathrm{cap} \times \sum_{j=1}^{t} x_j + \beta w_t \times \mathrm{cap} \times \sum_{j=1}^{t} y_j + z_t - s_t = d_t, \forall t \tag{2}$$

$$(p_s + pm_s)(1+\alpha)^t x_t + (p_w + pm_w)(1+\alpha)^t y_t + c_t(z_t + ss) - w \times s_t \leq I_t, \forall t \tag{3}$$

$$\sum_{t=1}^{T} fs \times x_t \leq F_s \tag{4}$$

$$\sum_{t=1}^{T} fw \times y_t \leq F_w \tag{5}$$

$$s_t \leq s_{\max} \times u_t, \forall t \tag{6}$$

$$z_t \leq z_{\max} \left(1 - u_t\right), \forall t \tag{7}$$

$$\frac{d_t - \beta s_t \times \text{cap} \times \sum_{j=1}^{t} x_j - \beta w_t \times \text{cap} \times \sum_{j=1}^{t} y_j + s_t}{d_t} \leq (1 - r_t), \forall t \tag{8}$$

$$x_t, y_t \geq 0, \text{ integer}, \forall t \tag{9}$$

$$z_t, s_t \geq 0, \forall t \tag{10}$$

$$u_t \in \{0, 1\}, \forall t. \tag{11}$$

The objective function is minimizing the cost of supplying energy from renewable sources and grid. The two first expressions calculate the investment and maintenance cost of the wind and solar energy sources by considering interest rate and inflation rate. The third expression shows the cost of buying energy from the grid, and the last one expresses hospital's income from selling energy to the grid.

Since supplying reliable energy is essential in hospitals, all the demand should be met which is presented in constraint (2). Constraint (3) represents the limitation of hospital's budget in each time slot. Area limitation for installing solar panels and wind turbines have been avouched by constraints (4) and (5), respectively. In each time slot, it is not possible to buy energy from the grid and sell energy to the grid simultaneously, which have been determined in constraints (6) and (7). Constraint (8) represents renewable energy sources' reliability. In this equation, the left expression shows the amount of energy which has been purchased from the grid which should be less than a specific amount. Constraints (9)–(11) indicates variable types.

3 Numerical Example

To show the performance of the model, it is solved for a hospital in Iran. The numerical example is carried out by using IBM ILOG CPLEX_Optimizer v12.3. As mentioned, the goal of this model is to find the optimal size of renewable energy sources (RESs) installed in the hospital during the ten years time horizon. For reducing forecasting errors, years are divided into months time slots (120 time slots). Hospital's monthly demand and budget are estimated during meetings with its experts. Some of the data used in the model are shown in Table 1.

The results are shown in Fig. 1. With scrutiny, it is obvious that investment is higher in first time slots. It is reasonable due to the positive inflation rate. Moreover, investments in solar panels are more than wind turbines owing to the fact that they have more productivity in the aforementioned hospital, and also they need fewer investment costs.

Table 1. Data related to time slots

Time	d(kW)	I (dollars)	βs	βw
1	12000	50000	0.5	0.23
2	12000	50000	0.6	0.3
⋮	⋮	⋮	⋮	⋮
120	15000	8750	0.25	0.3

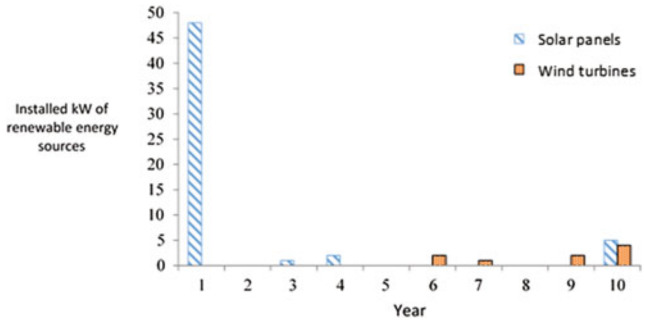


Fig. 1. Installed kW of wind turbines and solar panels in each year

Total solar panels which will be installed during ten years is 56 kW and installed wind turbines are 9 kW.

Sensitivity analysis is done for effective parameters on model such as inflation rate, interest rate, demand, reliability and etc. Some of them which are more important are elaborated in this paper.

In the aforementioned hospital, a high amount of budget is allocated to the first time slot because the decision for investment is given in this time slot. It is assumed that the budget is allocated to the hospital in the sixth time slot or at the beginning of the second year (13th time slot) with considering its interest. According to results, which is showed in Table 2, if the interest rate is between 15 and 20%, it is better to allocate budget in the first time slot. If the interest rate is between 20 and 24%, it is better to allocate budget in the 6th time slot.

Table 2. Inflation rate sensitivity analysis

Capital allocating method		Capital allocation in the first period	Capital allocation by considering its profit in the sixth month	Capital allocation by considering its profit in the second year
Objective function	Inflation rate = 15%	80226	82562	83379
	Inflation rate = 20%	67324	64824	67849

Table 3. Purchased price sensitivity analysis

	40% decrease in price	20% decrease in price	Current price	20% increase in price	40% increase in price	250% increase in price
Solar panels kW	56	56	56	56	56	56
Wind turbines kW	9	9	9	9	9	9
Amount of purchased energy from the grid	96460	90828	89744	88544	87416	53020

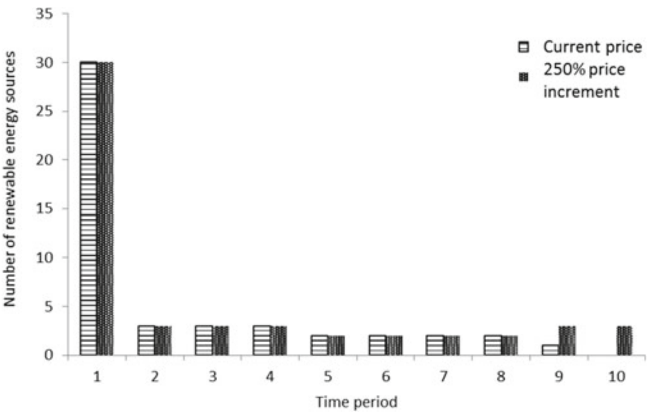


Fig. 2. Installed kW of wind turbines and solar panels in each year

When purchased price is increased, it is expected that investment in RESs is increased too. However, other scenarios can be happened due to other parameters. For example, it is possible that the number of RESs is decreased or even stay stable, but the time of investment is changing in these scenarios. According to results, shown in Table 3, with increment or decrease in purchased price, number of installed RESs is stable, and the time of investment is just changed. For understanding such changes, variations of investment time for current price and 250% price increment are shown in Fig. 2 for ten time slots. With current price, investment in wind turbines is started from 66th time slot; however, if purchased energy increases 250%, investment in wind turbines starts from 9th time slot, which causes to take advantage of wind generation in 57 more time slots.

Energy price is low in Iran; however, in developed countries such price is a significant amount. As a result, we compared current price in Iran with current price in United States, which is about 250% higher than energy price in Iran.

Table 4. Wind turbine utilization percentage sensitivity analysis

	30% increase in wind turbines utilization percentage	20% wind turbines utilization percentage	10% wind turbines utilization percentage	Current utilization percentage	10% wind turbines utilization percentage	20% wind turbines utilization percentage	30% wind turbines utilization percentage
Solar kW panels	48	48	51	56	57	60	61
Wind turbines kW	12	13	12	9	9	7	7

Table 5. Wind turbine utilization percentage sensitivity analysis

	20% increase in solar panels utilization percentage	10% increase in solar panels utilization percentage	Current utilization percentage	10% decrease in solar panels utilization percentage	20% decrease in solar panels utilization percentage
Solar panels kW	52	53	56	57	64
Wind turbines kW	4	7	9	13	13

Table 4 shows sensitivity analysis for wind turbine utilization percentage. When wind turbine utilization percentage increases, investment in solar panels decreases, and investment in wind turbines increases. Whatever the wind turbine utilization percentage increases, there is less need for wind turbines. With wind turbine utilization percentage reduction, the number of solar panels increases.

Table 5 presents sensitivity analysis for solar panels utilization percentage. By increment in solar panels utilization percentage, amount of installed kW of wind turbines decrease. Moreover, there is less need to install solar panels owing to the fact that more demand will be met by less solar panels with higher utilization percentage. On the other hand, by decrease in solar panels utilization percentage, the number of wind turbines increases as far as area constraint permits.

4 Conclusion

In this paper, long term scheduling model was presented for determining optimal sizing of renewable energy sources in hospitals. The model aimed at minimizing costs while maintaining reliability according to specific constraints in hospitals such as area limitation and budget constraint. By solving the model with real data of the hospital in Iran, the model can reduce 10% of costs comparing with the time when energy is only purchased from the grid. In presented model, planning horizon was divided into monthly time slots which caused an increment in forecasting errors. For further researches, authors can deal with daily or weekly data. Also in this model, the price of required field for installing wind turbines and solar panels was not considered. Authors can investigate the effect of field price on model outputs.

Acknowledgements. The authors would like to acknowledge Professor Mitsuo Gen for his scientific support.

References

1. Abd-El-Motaleb AM, Bekdach SK (2016) Optimal sizing of distributed generation considering uncertainties in a hybrid power system. Int J Electr Power Energy Sys 82:179–188
2. Atia R, Yamada N (2016) Sizing and analysis of renewable energy and battery systems in residential microgrids. IEEE Trans Smart Grid 7(3):1–10
3. Bizzarri G, Morini GL (2006) New technologies for an effective energy retrofit of hospitals. Appl Therm Eng 26(2–3):161–169
4. Mavrotas G, Diakoulaki D et al (2008) A mathematical programming framework for energy planning in services' sector buildings under uncertainty in load demand: the case of a hospital in athens. Energy Policy 36(7):2415–2429
5. Moradi MH, Eskandari M (2014) A hybrid method for simultaneous optimization of dg capacity and operational strategy in microgrids considering uncertainty in electricity price forecasting. Int J Electr Power Energy Syst 56(7):241–258
6. Pagliarini G, Corradi C, Rainieri S (2012) Hospital chcp system optimization assisted by trnsys building energy simulation tool. Appl Therm Eng 44(6):150–158
7. Renedo C, Ortiz A et al (2006) Study of different cogeneration alternatives for a spanish hospital center. Energy Build 38(5):484–490
8. Sharafi M, Elmekkawy TY (2014) Multi-objective optimal design of hybrid renewable energy systems using pso-simulation based approach. Renew Energy 68(68):67–79
9. Theo WL, Lim JS et al (2016) An milp model for cost-optimal planning of an on-grid hybrid power system for an eco-industrial park. Energy 116:1423–1441
10. Thomaidis NS, Santos-Alamillos FJ et al (2016) Optimal management of wind and solar energy resources. Comput Oper Res 66:284–291
11. Vaziri SM, Rezaee B, Monirian MA (2017) Bi-objective integer programming of hospitals under dynamic electricity price. In: Proceedings of the tenth international conference on management science and engineering management. Springer, pp 409–417
12. Wang YC, Mcpherson K et al (2011) Health and economic burden of the projected obesity trends in the USA and the UK. Lancet 378(9793):815–825
13. Yoshida S, Ito K, Yokoyama R (2007) Sensitivity analysis in structure optimization of energy supply systems for a hospital. Energy Convers Manag 48(11):2836–2843

Life Cycle Assessment of Waste Mobile Phone Recycling–A Case Study in China

Tingting Liu, Moudi Mahdi, and Liming Yao(✉)

Uncertainty Decision-Making Laboratory, Sichuan University, Chengdu 610064, People's Republic of China
lmyao@scu.edu.cn

Abstract. With the rapid pace of technological development, electronic waste (e-waste) has become one of the fastest growing waste streams in the world. The exponential growth of e-waste, which is disposed of in landfills, incinerated, or dismantled, now significantly contributes to the massive increase in overall waste. This paper assesses the waste produced from recycling mobile phones in China and examines the influence of the waste that is released into the environment. The mobile phone recycling process at formal recycling facilities in China is investigated as a case study. The results highlight the toxicity of the waste and the influence such waste has had on the environment. Life cycle assessment (LCA) was employed to analyze the mobile phone recycling process and Gabi was used to quantitatively indicate the flows and environmental impacts. The results showed that resources recovery, human health, and ecosystem quality are the main effects of the mobile phone recycling waste produced. Based on this quantitative analysis, to reduce the environmental impact of mobile phone recycling, the demand for recycled materials, environmental awareness, law enforcement, and the e-waste recycling system were found to be significant drivers.

Keywords: Waste mobile phone · Life cycle assessment · Recycling

1 Introduction

Electronic waste, also known as e-waste, has been one of the major contributors to the waste stream since the rapid growth of advanced technological products [6]. E-waste includes waste electrical and electronic products such as televisions, desktop computers, microwaves and mobile phones. In recent years, economic and technological development in China has seen as massive growth in mobile phone penetration and a consequent rise in mobile phone waste.

Currently over 40 million tonnes of e-waste is generated annually in the world and is growing every year. Of all the kinds of e-waste, the growth in mobile phone waste has been three times faster than total annual waste [19]. Current mobile phone using data indicates that there are 7.2 billion active mobile phones in 2016, with over 1.305 billion of these being in China, a penetration rate of 93.1%

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_113

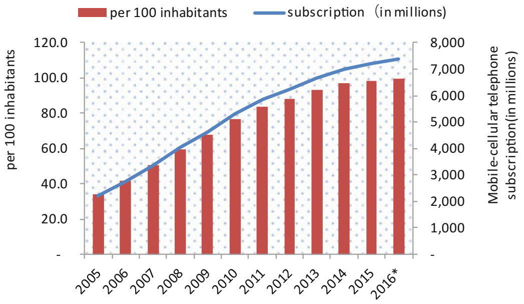


Fig. 1. Global mobile-cellular subscriptions, total and per 100 inhabitants, 2005–2016. Note: * Estimate

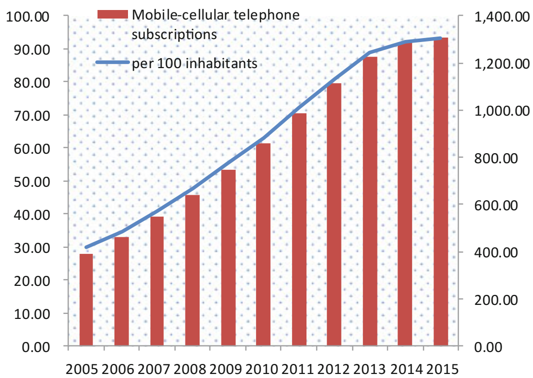


Fig. 2. China mobile-cellular subscriptions, total and per 100 inhabitants, 2005–2015

(see Figs. 1 and 2) [25]. About 77 million mobile phones are discarded every year in China, a figure the United Nations environmental program has predicted to be 7 times higher by 2020 [11].

The growing quantity of e-waste has begun to have a severe impact on the environment; especially in developing countries. As the leading manufacturing country of the world, China has become an "electronic waste treatment plant", with informal mobile phone recycling having become a major source of toxic pollution [13]. With rising climate change and water and land pollution concerns, mobile phone recycling waste has attracted increasing attention, forcing the mobile phone production industry to find solutions to decrease the environmental impact. In addition, various methods have been developed for mobile phone recycling to reduce the environmental damage. There has been a growing body of research into the recycling of waste mobile phones in China and the inherent problems. Wang et al. evaluated the potential yield of indium recycled from waste mobile phones in China [21], and Yi-Bo et al. and several others have examined the recovery and recycling of old mobile phones in China [23], and the other scholars did [4].

The main methods that have been applied to waste management analysis are data envelopment analysis(DEA) [2], the index decomposition analysis method(IDA) [12], and the structure decomposition analysis method (SDA) [1]. Life Cycle Assessment (LCA) was first used in 1969 when the Midwest Research Institute tracked and quantitatively analyzed the complete process of beverage containers from raw materials extraction to final disposal [3]. With development, the LCA has been included in the ISO14000 series of environmental management standards and has become an important support tool for international environmental management and product design [8]. LCA has been widely used in environmental assessments in such areas as food waste management [14], waste water [18], and plastic production as well as in waste electrical and electronic equipment (WEEE) management [10].

As an environmental management tool, LCA has expanded the boundaries of research systems, as it encompasses the entire life cycle of resource use, energy consumption and waste discharge and provides an evaluation of the potential environmental impact at each stage. This paper assesses the recycling of waste mobile phones in China using LCA. To date, analyses of the environmental impact of recycling waste mobile phones has lacked a systematic and scientific focus and has tended to ignore the management of discarded mobile phones [15]. In this paper, the life cycle assessment of the discarded mobile phone recycling process is quantitatively evaluated to identify the main factors influencing the environment and provide decision support for the waste management of other small electronic devices.

The remainder of this paper is structured as follows. In Sect. 2, the current recycling process in China is outlined, an abandoned mobile phone path analysis is conducted, and the latest recycling process explained. Section 3 applies the LCA to the mobile phone recycling process and gives the results. Section 4 gives suggestions and recommendations based on the analyses.

2 Mobile Phone Recycling in China

2.1 Recycling Status

China has had a rapid growth in e-waste streams over the past ten years [11], with mobile phone waste being the most rapidly growing. Because there has not been a regimented recycling system established in China, most waste mobile phones are kept, donated to others or discarded with other house-hold waste; the formal recovery rate has been only about 1% [25]. The recycling of batteries and other waste mobile phones parts has mainly been conducted through nonprofessional channels such as street vendors or small electronics shops or through recycling platforms on the internet. After recovery, refurbished mobile phones are generally sold in a second-hand market or sold at a low price in remote towns and villages. Mobile phones that cannot be refurbished often end up in illegal dismantling workshops, where workers extract the copper, gold, silver and other valuable metals using such processes such as acid leaching and open burning, processes which not only damage the health of workers but also severely pollute the surrounding environment.

From a product to waste, a mobile phone can change hands hundreds of times; they can be retained by consumers, returned to a retailer, sold to a second-hand shop or sent to recycling. The specific circulation paths are shown in Fig. 3.

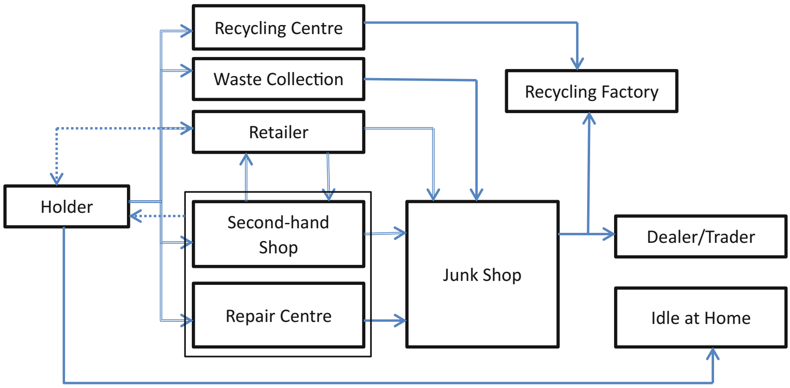


Fig. 3. The flow of waste mobile phones in China

2.2 Recycling Mode

This paper evaluates the pollution produced in the recycling phase from the perspective of material flows using the standard life cycle assessment method, examines the waste mobile phone recycling process in formal recycling facilities in China as a practical case study (see Fig. 4) and identifies the factors that adversely impact the environment in the recycling phase.

The brief description of the recycling process is as follows: after being transported to the waste treatment plant, the mobile phones are manually disassembled and the phone casings, printed circuit boards, lithium batteries, LCD screens and other main components extracted. Then, each of these components is processed separately. The ABS/PC plastic casings are crushed into plastic pellets then reused as raw materials, the stainless steel, copper and other metals are melted, with a further smelting processing step needed for the printed circuit boards and lithium ion batteries (recovery of electrolyte solvent) to recover the copper, gold, silver, palladium, cobalt, lithium, tin and other metals. LCD displays are generally incinerated to reduce environmental waste and the remaining residues are put into landfill.

Mentioned in the process: (1) the life cycle of the mobile phone at each stage from production, use, and disposal including the transportation of the various materials; however, the transportation and landfill residue stage after disassembly are not taken into account. (2) only lithium ion scrap mobile phone batteries and mobile phone LCD displays are considered.

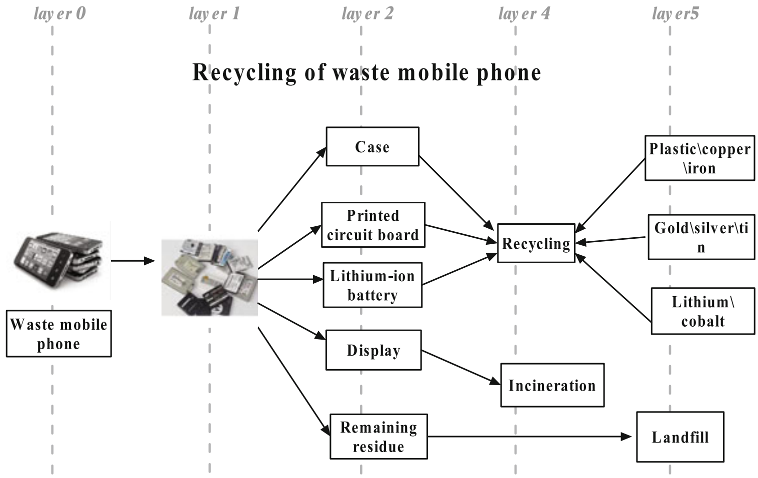


Fig. 4. Recycling process

3 LCA on Recycling Waste Mobile Phone

The LCA has four interrelated steps, which we applied to the mobile phone recycling process, the rough contents for which are shown in Fig. 5.

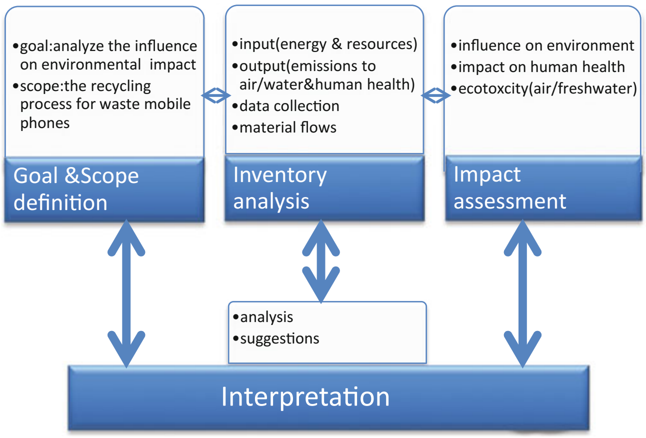


Fig. 5. Detailed content of the four steps

3.1 Goal and Scope Definition

The study object in this paper is the abandoned mobile phone. The functional unit is set to a tonne of phones of various brands and models, with the data for each component being set at an average value. The study scope for the system

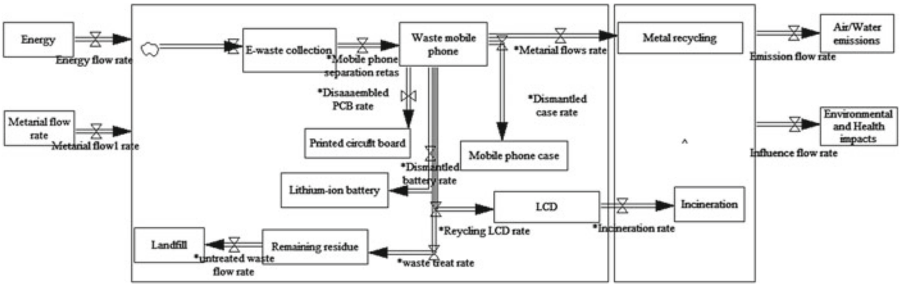


Fig. 6. Flows at a formal recycling facility

is set for disassembly, waste collection, sorting, processing and final disposal (Fig. 6) and an inventory analysis and impact assessment for the energy and resource consumption and pollutant emissions at each stage is established.

3.2 Inventory Analysis

From relevant research data regarding domestic and foreign waste mobile phone processing and combined with the Chinese current waste recycling, the main process inventory data related to the recycling process was determined (see Table 1). The data in this paper was taken from the ecoinvent database [7];

Table 1. Inventory analysis

Process flow	Inventory data
Disassembly	GLO: manual treatment plant, WEEE scrap [Recycling] = 2500 tone/yr. GLO: mechanical treatment plant, WEEE scrap [Recycling]/GIO/IU = 50,000 tone/yr
Plastic recycling	GLO: manual treatment plant, WEEE scrap [Recycling] = 2500 tone/yr
Iron metal	RER: steel, electric, un- and low-alloyed, at plant [Beneficiation]
Metal	SE: copper, secondary, from electronic and electric scrap recycling, at refinery [Beneficiation]
Printed Circuit Boards (PCBs)	GLO: disposal, treatment of printed wiring boards [Recycling] SE: gold, secondary, at precious metal refinery [Beneficiation] SE: silver, secondary, at precious metal refinery [Beneficiation] SE: palladium, secondary, at precious metal refinery [Beneficiation]
Tin	RER: tin, at regional storage [Beneficiation]
Li-ions batteries	GLO: disposal, Li-ions batteries, hydrometallurgical [Recycling] GLO: cobalt, at plant [Beneficiation] GLO: lithium hydroxide, at plant [inorganics]
LCD	CH: disposal, LCD module, to municipal waste incineration [municipal incineration]
Landfill	Disposal, municipal solid waste, 22.9% water, to sanitary landfill/CH U

as an overview and methodological framework for the LCA, the ecoinvent database is the most consistent and transparent life cycle inventory database in the world [22].

Inventory analysis involves the tracking of the input and output flows of a product system including the materials, water and energy used and the waste released to the air, land and water [17]. In the model developed in this paper, energy and raw materials consumption is taken into account and GaBi, a powerful life-cycle assessment tool, is used to generate the model from the inventory data to assess the sustainability solutions [24].

Fig. 7. Gabi built model

As shown in Fig. 7, the input for this model is 1000 kg of waste mobile phones.

3.3 Life Cycle Impact Assessment

LCIA is a qualitative and quantitative analysis of environmental impacts based on the elements listed in the inventory analysis (LCI) [20]. It is used to evaluate and interpret the environment impacts of a product system by assigning quantifiable measurements and connecting the input and output material inventories to obtain the damage points of a specified boundary system.

Generally speaking, the LCIA includes the following steps: classification of the elements listed in the inventory analysis process; a qualitative and quantitative analysis of the elements; identification of the major environmental factors in all aspects of the system; and analysis and judgment of the environmental factors. This paper identifies and analyzes the major environmental factors in the mobile phone recycling process. The environmental factors include ecotoxcity, cancer effects and human ecotoxicity [16].

3.4 Interpretation

In this section, the evaluation and improvements for the mobile phone recycling process are discussed, after which, with the aim of reducing the burden on the environment, suggestions are given to improve the current recycling mode of

China. In line with the goals and scope of the analysis, the recycling process results were evaluated, from which we were able to come to conclusions and develop recommendations.

The inventory results from the model focused on three main aspects. The first was the recovery of resources and materials. Valuable materials, such as copper, gold, and silver, which can be recycled and reused, were retrieved and the other materials treated as waste. The second aspect was human health concerns. Many related environmental issues such as carcinogenic materials, climate change, ozone depletion, radiation, and organic and inorganic respiratory irritants can cause various diseases, premature death and other non-normal deaths, reducing life expectancy. This aspect was measured in units of disability adjusted life years (DALY), which examined two kinds of health losses; disability and death; to determine the social impact of possible medical conditions on population health. This indicator has also been adopted by the World Bank and the World Health Organization. The third aspect focuses on ecosystem quality which refers to the ecological environmental biodiversity loss within a given time and space caused by environmental issues such as acidification/eutrophication of land and water bodies, toxicity, land-use changes [5], and air emission [9]. The measurement units for this aspect were the potentially disappeared fraction (PDF) m^2yr and the potentially affected fraction (PAF) m^2yr.

4 Discussion

The Chinese recycling waste mobile phone process was examined in this paper using LCA as the evaluation tool. A model was proposed and the possible three impact factors; resources, human health, and ecosystem quality; were analyzed to assess the environmental impact of the recycling process. In the model building procedure, resources consumption and material emissions were found to have a certain degree of impact on the environment, and further research can extract specific data to analyze the effects of specific substances such as heavy metals, CO_2 and NO_2. Because of the speed of mobile phone technological development, the rate of mobile phone disposal is expected to increase, further impacting the environment. If waste mobile phones can be effectively recycled, there will be less environmental and health effects, improving both the economy and the residents quality of life. This paper gave an assessment on the recycling process of waste mobile phone in China, identified the main factors affecting the environment. On the basis of this analysis, proposals were given to improve the eco-efficiency in recycling process. First, relevant policies should be presented to standardize the recycling behavior of citizens and market. Second, the recycling technology should be improved. Besides, a more eco-efficient recycling model is needed. The three main factors in this paper; resources, human health, and ecosystem quality; will be taken into account in the model.

Acknowledgements. The work is supported by the National Natural Science Foundation for Young Scholars of China (Grant No. 71301109), and Soft Science Program of Sichuan Province (Grant No. 2017ZR0154).

References

1. Ang BW, Xu XY, Su B (2015) Multi-country comparisons of energy performance: the index decomposition analysis approach. Energy Econ 47:68–76
2. Barba-Gutiérrez Y, Adenso-Díaz B, Lozano S (2009) Eco-efficiency of electric and electronic appliances: a data envelopment analysis (DEA). Environ Model Assess 14(4):439–447
3. Beccali M, Cellura M, Finocchiaro P (2014) Life cycle assessment performance comparison of small solar thermal cooling systems with conventional plants assisted with photovoltaics. Energy Procedia 104(1):893–903
4. Bian J, Bai H, Li W (2016) Comparative environmental life cycle assessment of waste mobile phone recycling in China. J Clean Prod 131:209–218
5. Bieda B (2008) The use of life cycle assessment (LCA) conception for arcelormittal steel poland S.A. energy generating-KRAKW plant case study. Wydawnictwo Politechniki Krakowskiej Im Tadeusza Kociuszki
6. Buekens A, Yang J (2014) Recycling of weee plastics: a review. J Mater Cycles Waste Manag 16(3):415–434
7. Domain TE (2016) Life cycle inventory database. http://www.ecoinvent.org/
8. Guinee JB (2002) Handbook on life cycle assessment operational guide to the iso standards. Int J Life Cycle Assess 7(5):311–313
9. Gurjar BR, Butler TM, Lawrence M (2008) Evaluation of emissions and air quality in megacities. Atmos Environ 42(7):1593–1606
10. Harris R (2015) Life cycle assessment of post-consumer plastics production from waste electrical and electronic equipment (weee) treatment residues in a central european plastics recycling plant. Sci Total Environ 529:158–167
11. Herat S, Agamuthu P (2012) E-waste: a problem or an opportunity review of issues, challenges and solutions in Asian countries. Waste Manag Res J Int Solid Wastes Public Clean Assoc Iswa 30(11):1113–1129
12. Korica P, Cirman A, Gotvajn A (2016) Decomposition analysis of the waste generation and management in 30 European countries. Waste Manag Res
13. Liu X, Tanaka M, Matsui Y (2006) Electrical and electronic waste management in China: progress and the barriers to overcome. Waste Manag Res J Int Solid Wastes Public Clean Assoc Iswa 24(1):92–101
14. Lundie S, Peters GM (2005) Life cycle assessment of food waste management options. J Clean Prod 13(3):275–286
15. Middendorf A (2008) Assessment of toxicity potential of metallic elements in discarded electronics. J Environ Sci 20(11):1403–1408
16. Navarro E, Baun A, Behra R (2008) Environmental behavior and ecotoxicity of engineered nanoparticles to algae, plants, and fungi. Ecotoxicology 17(5):372–86
17. Organization IS, Iso ISO (2006) Environmental management-life cycle assessment-principles and framework. International Standard ISO
18. Pradel M, Aissani L, Villot J (2015) From waste to product: a paradigm shift in LCA applied to wastewater sewage sludge. In: Setac Europe meeting
19. Ramesh BB, Parande AK, Ahmed BC (2007) Electrical and electronic waste: a global environmental problem. Waste Manag Res J Int Solid Wastes Public Clean Assoc Iswa 25(4):307–318
20. Russo G, Vivaldi GA, Gennaro BD (2015) Environmental sustainability of different soil management techniques in a high-density olive orchard. J Clean Prod 107:498–508

21. Wang HG, Wu YF et al (2016) An evaluation of potential yield of indium recycled from waste mobile phone in China. In: Materials science forum
22. Wernet G, Bauer C, Steubing Bea (2016) The ecoinvent database version 3 (part i): overview and methodology. Int J Life Cycle Assess:1–13
23. He YB, Li LQ (2007) A study on recovery and recycle of old mobile phones in China. Nonferrous Met Eng Res
24. Zhang H (2015) Product sustainability software. http://www.gabi-software.com/
25. Zhao H (2016) ITU World Telecommunication /ICT Indicators database. http://www.itu.int/

A Malmquist Index-Based Dynamic Industrial Green Efficiency Evaluation in Sichuan Province

Xuhong Liu(✉) and Xiaowen Jie

Business School, Sichuan University, Chengdu 610064, People's Republic of China
liuxh_99@163.com

Abstract. In this paper, the DEA-Malmquist index is used to evaluate the industrial green efficiency of 18 cities in Sichuan province during 2008–2014. The study found that in the overall level, the industrial green efficiency of Sichuan prefecture level cities' is low, but it is growing year by year in recent years. The technical efficiency index grows rapidly, and the technological progress index weakened the role of technical efficiency. Industrial green efficiency of Chengdu Plain economic zone is the highest in the four economic zones, followed by the Panxi economic zone, South Sichuan economic zone, lowest in Northeast Sichuan economic zone. Industries need to increase the intensity of the green transformation in Sichuan province.

Keywords: Malmquist index · Dynamic evaluation · Industrial green efficiency · Sichuan province

1 Introduction

The industrial developments of China have received great achievement recent years, but the past extensive industrial mode was at the cost of environment pollution and excessive resources consumption, which results in the problems of resource constraint and environment pressure more visibly. And the resource utilization has also become one important factor to measure the competitiveness of a country's manufacturing industry, and the industrial green development is the necessary way to promote international competitiveness. In order to change the industrial mode of high involvement, high consumption and high pollution, to promote the region industrial economic development, the 13th Five Year Plan outline of national economic and social development had put forward higher requirement for industrial green development.

Sichuan province, as the most important economic area of China western, its industrial green development is in what level, whether the industrial green efficiency has improved. The deep research of these issues would be great theoretical and practical significance. So based on the urgent need of green development, the presented paper takes the prefecture-level city as research objects, by using DEA research method to dynamic evaluate the green efficiency of industrial inputs and its output. And propose countermeasures and suggestions to improve industrial green efficiency of Sichuan province.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_114

2 Literature Review

Industrial green development is the inner requirement and inevitable choice of new industrialization. United Nations Industrial Development Organization (UNID) definition that green development is a sustainable industrialization mode of the production and consumption by effective utilization energy resources, low carbon emission, low waste reduction, during the industrialization development process to expand in order to eliminate poverty and create employment opportunities. The proposed research consider that the industrial green development is under the theory guidance of sustainable development and circular economy, through green design, green production, green management and other aspects of the control, to realize the high utilization rate of energy resource, low environmental disruption, high comprehensive benefits development mode.

At present, the foreign scholars are focused to research the connotation of industrial green development and evaluation of its efficiency. Eiadat et al. [2] and Nagesha [11] selected patent number of green process and the carbon dioxide emissions to assess the industrial green efficiency, Honma and Hu [6] used the random front model to evaluate the industrial green efficiency. There are scholars had researched the influence factors of green development on the specific industry, Vasauskaite et al. [15] put forward R&D investment, energy saving technology or energy management system can promote the sustainable development of Lithuania furniture industry. Fleiter et al. [4] evaluated the 17 process technologies of the pulp and paper industry in Germany, and considered that the innovation of paper mills and paper drying technologies in heat recovery systems is the most influential technology for energy conservation and green development. Mukherjee et al. [9] proposed computer-aided process design can optimize the energy and material flow to minimize the impacts from other harmful emissions. Kanagaraj et al. [10] reviewed different cleaner technological methods of reducing the generation of solid and liquid wastes in leather industry.

Domestic scholars Chen et al. [1] researched the influencing factors of industrial green development. Wang and Huang [16] according to the five indexes of industrial and social development, recycling, resource reduction, pollution reduction, resource and environmental security, used AHP method to evaluate the development of industrial circular economy in Jiangsu province from 1985 to 2003. And they introduced the concept of "obstacle degree", to quantitatively diagnosis obstacle in different stages of industrial circular economy development in Jiangsu province quantitative. Tu and Xiao [14] established frontier function model of environmental by directional distance function, and calculated the green total factor productivity of industrial enterprises above designated size, with energy consumption and environmental pollution in 30 provinces of China in 1998–2005. As well the pollutants considered the emissions of industrial sulfur dioxide. It has found that the industrial development In the eastern coastal areas was more harmonious with the environment, and the environmental technology efficiency was generally low in the central and western regions. Mei et al. [8] established the evaluation index system of green development efficiency of coal industry by improved DEA. Wu [17] analyzed China's industrial total factor productivity (TFP) growth in 1998–2007 under environmental constraints through Malmquist-Luenberger index. The results showed that the TFP of western industry was growing fastest, eastern followed, while the central growth was the slowest. Zhang and Zhu [22] took the non-expected industrial exhaust emissions as input variables, and measured the efficiency of industrial innovation based on green growth. It was considered that there was a phenomenon of increasing returns to scale in the Western region. Yang [19] used DEA method to measure the efficiency of low carbon economy in Sichuan Province, based on the analysis of energy consumption and carbon emissions, and concluded that the efficiency of developing low-carbon economy was still not high in Sichuan Province, which was mainly restricted by technology shortage and scale bottleneck.

Therefore, the researches of industrial productivity considering of resource environmental factors are basically from the regional perspective. There are few studies on the efficiency of industrial development from the perspective of Economic Zone. Or mostly concentrated in the eastern region of China, such as the Pearl River Delta urban agglomeration, Shandong peninsula urban agglomerations and other places, the city of the western region was less discussed. So the presented paper based on the original data of prefecture level cities in Sichuan Province, evaluates the status of its industrial green development and lays the foundation for further research.

3 General Situation of Industrial Green Development in Sichuan

At present, industrial development in Sichuan Province is in the middle stage, but according to the environmental Kuznets curve, the negative impact on energy and the environment is still relatively large of the current industrial development in Sichuan [7]. Moreover, Sichuan is a traditional resource and energy province, in which has proven reserves of 132 kinds of mineral resources, accounting for 70% of the total resources in China. This will inevitably lead to its industrial development mainly resource oriented. From 2005 to 2015, the average annual growth of GDP in Sichuan province is 8.1%, but the average annual growth of total energy consumption is 3.5%, which put enormous pressure on Sichuan's resources.

It can be seen from Fig. 1 that, energy consumption per unit of industrial added value and energy consumption per unit of GDP in Sichuan continued to decline (2005–2013), decreased by 48.64% and 32.67% respectively. This shows that the efficiency of industrial energy utilization in Sichuan province has increased rapidly, and the industrial structure of its low-end has improved. But on the other hand, by the end of 2014, raw material resources traditional industry in Sichuan province accounted for up to 43%, and low-end manufacturing industry accounted for nearly 30%. It indicates that Sichuan province still needs to further increasing the green structure transformation, so as to improve the utilization efficiency of resources and energy. Figure 2 showed that the discharge of industrial waste water, waste gas and solid waste (hereinafter referred to as three wastes) per unit of GDP in Sichuan Province decreased year by year during the period of 2005–2013, with the decreasing rate of 38.96%, 32.07% and 85.19%. These were higher than the decreasing rate of national industrial three wastes emissions (30.63%, 21.85% and 72.89%). It indicates that in the process of industrial development, Sichuan Province compared to the national, the level of mitigation to environmental pollution is higher. In addition, although the industrial waste water, waste gas per unit of GDP in Sichuan province is lower than the national level after 2010, the industrial solid waste generated per unit of GDP is still higher than the national recently. It showed that the situation of environmental pollution in Sichuan province is still grim, and it is an important way of strengthening the industrial green transformation to ease the pressure of resources and the environment.

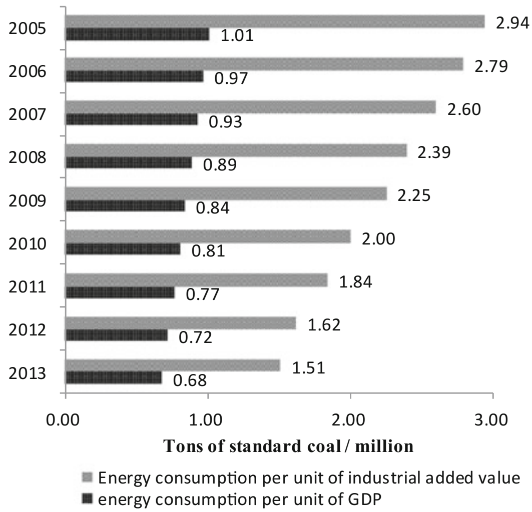


Fig. 1. Resource consumption situation of Sichuan province in 2005–2013 (Source: China Statistical Yearbook)

4 The Measurement of Industrial Green Efficiency

4.1 Malmquist Index Model

Malmquist index was proposed by the famous economist Malmquist in 1953, then Cavers et al. and Fare applied this index to estimate productivity. It can describe multiple input variables and multiple output variables of production technology, and is a method of measuring total factor productivity using panel data by

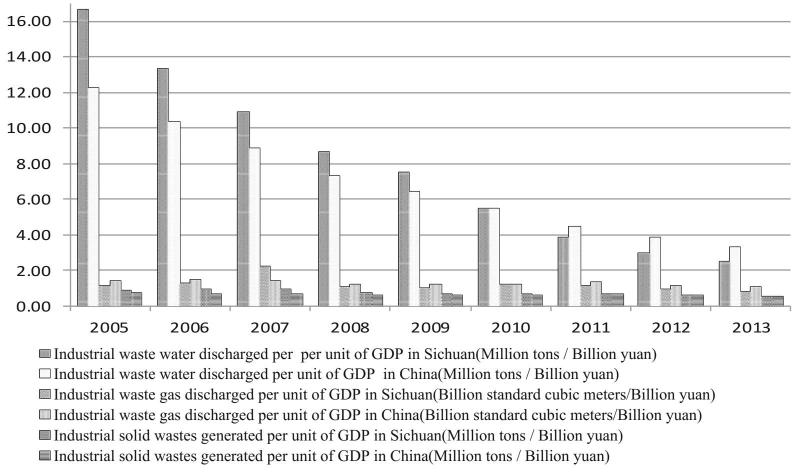


Fig. 2. Comparison of “three wastes” emissions in Sichuan province and China in 2005–2013 (Source: calculated according to the Chinese Statistical Yearbook 2006–2014)

nonparametric. In empirical analysis, researchers generally use Malmquist index based DEA by Fare et al. [3]. Its true value is approximately:

$$M_0{}^{t+1}\left(x^{t+1}, y^{t+1}, x^t, y^t\right) = \frac{D_0{}^{t+1}\left(x^{t+1}, y^{t+1}\right)}{D_0{}^{t}\left(x^{t}, y^{t}\right)} \times \sqrt{\frac{D_0{}^{t}\left(x^{t+1}, y^{t+1}\right)}{D_0{}^{t}\left(x^{t}, y^{t}\right)} \times \frac{D_0{}^{t+1}\left(x^{t+1}, y^{t+1}\right)}{D_0{}^{t+1}\left(x^{t}, y^{t}\right)}}. \tag{1}$$

In formula (1), M_0^{t+1} is the index for decision making unit DMU0 in t+1 period, $(x^t,\ y^t)$ and $(x^{t+1},\ y^{t+1})$ represent input-output vectors for period t to period $t+1$ respectively, D_0^t and D_0^{t+1} represent the distance function taking technology as a reference for period’s t and $t+1$ techniques. Further, under the assumption of variable returns to scale, formula (1) can be rewrited as follows:

$$\begin{aligned} M_0{}^{t+1}\left(x^{t+1}, y^{t+1}, x^t, y^t\right) = & \frac{D_0{}^{t+1}\left(x^{t+1}, y^{t+1}|V\right)}{D_0{}^{t}\left(x^{t}, y^{t}|V\right)} \\ & \times \left[\frac{D_0{}^{t+1}\left(x^{t+1}, y^{t+1}|C\right) D_0{}^{t}\left(x^{t}, y^{t}|V\right)}{D_0{}^{t}\left(x^{t}, y^{t}|C\right) D_0{}^{t+1}\left(x^{t+1}, y^{t+1}|V\right)}\right] \\ & \times \sqrt{\frac{D_0{}^{t}\left(x^{t+1}, y^{t+1}\right)}{D_0{}^{t+1}\left(x^{t+1}, y^{t+1}\right)} \times \frac{D_0{}^{t}\left(x^{t}, y^{t}\right)}{D_0{}^{t+1}\left(x^{t}, y^{t}\right)}}. \end{aligned} \tag{2}$$

In formula (2), V represents a variable size, C represents the same size. The first item represents the pure technical efficiency index (Pech), the second represents the scale efficiency index (Sech), and the third represents the technical progress index (Techch), that is Tfpch = Effch × Techch = Pech × Sech × Techch. Among them, the technical efficiency index measures the change degree to the production frontier of each decision unit from period t to period $t+1$, the technical progress index measures the moving degree of production frontier from period t to period $t+1$, the pure technical efficiency index reflects the technical efficiency when the scale returns unchanged, and the scale efficiency index reflects the optimal production scale of the decision unit. This paper uses directional output variables to measure the industrial green efficiency of prefectural-level cities in Sichuan. When the index is equal to 0, indicating it has no effect on the industrial green efficiency. If the index is greater than 1, indicating a promotion effect, or vice versa, showing there is a blocking effect.

4.2 Evaluation Indexes and Data Sources

The selection indexes in industrial green efficiency evaluation system are not only to meet the needs of industrial economic development, but also to meet the demands of environmental and social development. The presented paper selects evaluation index of industrial green efficiency combining with previous studies [13,18,21], and on the basis of following principles of scientific, systematic and operational. As shown in Table 1, input indicators select key elements of industrial development process. X1 and X2 represents human input and capital input of industrial development respectively. X3 reflects investment in energy application. Output indicators select total industrial output value (Y1) and total

Table 1. Evaluation index of industrial green efficiency

Indicator types	Index code and index name	Unit	Data sources
Input index	X1: employment number[a]	Ten thousand people	Statistical Yearbook of Sichuan Province
	X2: total investment in fixed assets[b]	Billion yuan	China City Statistical Yearbook
	X3: industrial added value of energy consumption	Tons of standard coal/million	
Output index	Y1: total industrial output value	Billion yuan	
	Y2: total discharge of industrial waste pollution[c]	Ton	

Notes Based on data availability. [a]Employment number selects the number of employed persons at the year-end in the mining, manufacturing and electricity, gas and water production and supply industries. [b]Total investment in fixed assets selects gross fixed assets of industrial enterprises above-scale on behalf of existing production and operation scale and technical level of industrial enterprises. [c]Because the comprehensive utilization rate of industrial solid waste in cities are all high and have no remarkable difference, so the output indicators mainly selects industrial emissions of waste water, waste gas (SO_2) and smoke (powder) dust.

discharge of industrial waste pollution (Y2) respectively reflects the industrial output and pollution emissions of each cities. In order to eliminate the influence of inflation, the total industrial output value of current year was converted to constant price in 2005, according to the ex-factory price index of industrial products. And the index unit of total discharge of industrial waste pollution was processed.

Each evaluation index data are from statistical yearbook of Sichuan and City Statistical Yearbook during 2005 to 2014. Some missing data has been processed accordingly to ensure the accuracy of indicators. Due to a large number of missing data in Ganzi, Aba, Liangshan area, this paper takes other 18 cities in Sichuan province as the research object.

Among them, total discharge of industrial waste pollution uses entropy weight method to calculate the index weight, and take reciprocal at last as a bad output. The process of calculating as follows:

(1) Standardizing the industrial emissions of waste water, waste gas (SO_2) and smoke (powder) dust. X_{ij} Indicates J index values in I area:

$$r_{ij} = \frac{(X_{j\,\max} - X_{ij})}{(X_{j\,\max} - X_{j\,\min})}.$$

(2) Calculating the entropy of each index, m is the number of evaluation units, n indicates the number of indicators:

$$H_i = -\frac{1}{\ln_m} \sum_{i=1}^{m} \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}} \times \ln \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}} \left(\text{while } \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}} = 0,\ \ln \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}} = 0. \right)$$

(3) Obtaining the entropy weight of every index. To keep the longitudinal comparability of the dynamic evaluation results, according to the method of Zhang and Zhou [23], the presented paper takes the average entropy weight as the final entropy weight to avoid the differences in each year:

$$W_j = \frac{1 - H_j}{n - \sum_{j=1}^{n} H_j}.$$

(4) Calculating the total discharge of industrial waste pollution (abbreviation to TIW):

$$TIW = \sum_{j=1}^{n} r_{ij} w_j.$$

5 Result

5.1 Overall Analysis of Industrial Green Efficiency in Sichuan

The paper using DEAP 2.1 software to calculate the Malmquist index of industrial green efficiency of Sichuan province and its decomposition during 2005–2014 as shown in Table 2. It can be seen that the level of Sichuan's industrial green efficiency is low, but the development trend is increasing year by year. This is consistent with Jia [7] results, which the Sichuan's green industrial development rate was still lagging behind, energy efficiency has gradually increased. But the energy pressure is still large, environmental pollution has been curbed but the situation is still grim. From the index decomposition terms, the impetus for growth of industrial green efficiency comes from technical efficiency, and its average annual growth rate is 2.1%. The improvement of technical efficiency is mainly due to pure technical efficiency and scale efficiency, which increased by 1.6% and 0.6% respectively. But the technical progress efficiency is low, far below the effective value of 1. It indicates that it should increase the R&D, introduction, digestion and absorption of green technology, to drive the green development of Sichuan industry. One of the highest levels of Sichuan industrial green development is in 2013–2014 with an increase of 16.6%, and mainly depends on the improvement of technological progress index. This may be due to the governments and enterprises had paid more attention to the green development of industry in recent years, as well as the lag effect of green technology has begun to produce.

Table 2. Average growth rate per annum and decomposition of TFP in Sichuan province (2005–2014)

Year	Effch	Techch	Pech	Sech	Tfpch
2008–2009	1.071	0.779	1.046	1.024	0.835
2009–2010	0.953	0.804	1.001	0.952	0.767
2010–2011	1.083	0.759	1.001	1.082	0.822
2011–2012	0.963	0.973	1.001	0.962	0.937
2012–2013	1.06	0.845	1.013	1.047	0.896
2013–2014	0.994	1.173	1.016	0.978	1.166
Mean	1.021	0.887	1.016	1.006	0.906

5.2 Analysis of Industrial Green Efficiency Among the Cities of Sichuan Province

According to the year data from 2005 to 2014, we further study the Malmquist index of industry green efficient and its decomposition terms item of Sichuan province; the results are shown in Table 3. It is easy inferred that there is big difference between different regions from the table date. The industry green efficient of Chengdu, Deyang, Mianyang, Yibin and Dazhou has increased, which accounts for 27.8%, Chengdu is 6% with maximal growth rate. The industry green efficient growth rate of rest of 13 regions is negative, Guangyuan is −20.5% with the lowest growth rate. This indicates that the industry prosperous of the latter is at the cost of taking up and consuming a large number of ecological resources. And the former is the beneficiary of development experience and mature technology learning.

Table 3. Average growth rate per annum and decomposition of TFP in cities (2005–2014)

Region		Effch	Techch	Pech	Sech	Tfpch
Chengdu Plain	Chengdu	1.054	1.006	1	1.054	1.06
	Deyang	1.055	0.974	1.081	0.976	1.028
	Mianyang	1.02	0.98	1.02	1	1
	Suining	1.025	0.864	1.056	0.971	0.886
	Leshan	1.055	0.829	1.011	1.044	0.875
	Meishan	1	0.822	1	1	0.822
	Yaan	1.042	0.843	1.054	0.989	0.879
	Ziyang	0.954	0.869	0.943	1.012	0.829
South Sichuan	Guangyuan	0.986	0.806	0.988	0.998	0.795
	Nanchong	1.064	0.868	1.056	1.008	0.924
	Guangan	1.009	0.842	1	1.009	0.85
	Dazhou	0.998	1.002	1	0.998	1
	Bazhong	0.987	0.807	0.987	1	0.797
Northeast Sichuan	Zigong	1.028	0.874	1.03	0.998	0.898
	Luzhou	1.005	0.867	1.003	1.002	0.871
	Neijiang	0.965	0.846	0.988	0.977	0.817
	Yibin	1.037	0.968	1.037	1	1.004
Panxi	Panzhihua	1.1	0.894	1.03	1.068	0.983
Sichuan	Mean	1.021	0.887	1.016	1.006	0.906

Note: Because of the missing statistical data of Liangshan state, Ganzi state and Aba state, this paper does not studied the West Sichuan Economic Zone. And considered the large industrial scale of Panzhihua, it was chosen to represent the overall level of Panxi Economic Zone.

Combined the Table 3 and Fig. 3, we analysis deeply the dynamic change trend and the distribute state of the Malmquist index resolution item in Sichuan province. Figure 3 reflects the growth pattern and characteristic of the industrial green efficiency. The cross-point of the horizontal-vertical coordinate is the equalization point of the technology efficient index and technology progress index, by which the 18 prefecture-level cities of Sichuan province are divided into 4 quartiles. In the first quartile, the Effch is higher and Techch is lower than the average. It indicates that these regions with lower technology contribution and simplex elements driven mode. The Effch and Techch are higher than equalization in the second quartile, these regions with reasonable element input scale, high technology import and absorption contribution rate. It promotes the growth of the

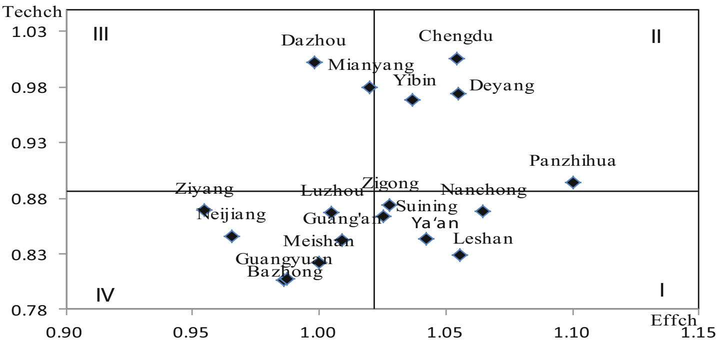


Fig. 3. The Techch and Effch distribution of the industrial green efficiency of 18 prefecture-level cities of Sichuan province

industrial green efficiency. The Techch is higher than Effch in the third quartile, it manifests that the technology level of these regions is above of production frontier. But it needs to improve the scale efficiency and pure technology efficiency. The Effch and Techch are lower than province equalization in the fourth quartile, these regions with low industrial green level and weak sustainable development due to the technical constraints and scale bottlenecks.

In the meantime, from the four major economic regions data of Chengdu Plain of Sichuan, Northeast Sichuan, South Sichuan, and Panxi (the presented paper have not studied the economic region of northwest of Sichuan due to the large missing data), it indicates that the industrial green efficiency is a gradual reduce step distribute from north to south, and from west to east, as show in Table 4.

The industrial green efficiency of Panxi head the list, secondly are Chengdu Plain of Sichuan and South Sichuan, the worst is Northeast Sichuan, its year equalization reduce is 12.7%. The technology progress indexes are inefficiency state of the four economic regions, from Malmquist index and its resolution items. Besides the south of Sichuan, the technology efficiency and its resolution items of the other three economic regions are progressing, which reveals that its resource allocation ability, resource utilization rate, and scale efficiency are valid.

The driving force of technical efficiency in Chengdu Plain Economic Zone is mainly pushed forward by pure technical efficiency. This also shows that the radiation driven role of Chengdu, as the capital of Sichuan Province, has significant advantages in terms of production factors, management level and system construction. The improvement of technical efficiency of Panxi Economic Zone mainly originates in its advantage of scale efficiency (an increase of 6.8%). It also shows that the resource input and its utilization efficiency have achieved scale benefit. Conversely, the scale efficiency of South Sichuan Economic Zone has decreased by 0.6%, which may be related to the unreasonable input structure of production factors.

Table 4. Average growth rate per annum and decomposition of TFP in economic zone (2005–2014)

Region	Effch	Techch	Pech	Sech	Tfpch
Chengdu Plain Economic Zone	1.022	0.883	1.023	1	0.903
Northeast Sichuan Economic Zone	1.009	0.865	1.006	1.003	0.873
South Sichuan Economic Zone	1.009	0.889	1.015	0.994	0.898
Panxi Economic Zone	1.1	0.894	1.03	1.068	0.983

6 Conclusions and Suggestion

Combined with the existing problems and deficiencies in the process of industrial green development in Sichuan province, there are several measures could be taken.

(1) Adjusting the industrial structure within industry.

To eliminate backward production capacity with high energy consumption and low added value, and reduce the industries proportion of high energy consumption and high pollution [8], such as steel, cement and others. Developing renewable energy sources actively, like wind, solar, bio-energy and others to improve coal-based industrial structure. Besides, it can expand the scale of environmental protection industry, for instance, solid waste treatment industry and sewage treatment industry.

(2) Promoting technological innovation.

It is beneficial to reform and upgrade traditional industries, to strengthen technological transformation, thus promoting the development of green recycling economy. Positively implementing environmentally friendly technologies and industrial process, such as industrial energy and water saving, clean production, recycling and others, to enhance green technology capability [12]. Encouraging university–industry collaboration to promote the industrialization of green technology. Protecting intellectual property is to enhance the enthusiasm of enterprises to develop green technology, and solve the market barriers. The government can set up innovation platform with the opportunities of the eastern industry transfer to the west, to enhance the ability of independent innovation.

(3) Adjusting scale structure to improve industrial green efficiency.

Using "incentive" environmental regulation tools based on market tools, and establish market based emission regulations and industry dynamics to promote industrial scale efficiency [5]. Increasing the environmental surveillance of medium-sized enterprises and implementing environmental measures with powerful binding. In the aspects of the industrial development policy, it's should play an exemplary role of leading enterprise, encourage the joint development of small and medium enterprises, enhance the linkage with large enterprises in the upstream and downstream industry chain. In addition, establishing green network supply chain to reduce the damage to ecological environment and achieve the goal of improving industrial green efficiency.

(4) Accelerating the construction of green manufacturing system.

Firstly is to improve “green performance” in product design, manufacturing, transportation, application and other product life cycle. Secondly is to create and promote green factories and green areas, with high technology content, high resource utilization and low environmental pollution. Thirdly is to establish green manufacturing standards and clean energy policies [20], which can evaluate the resource utilization and the cost of social environmental, thus to constraint and guide industrial green development.

The presented paper studies the industrial green efficiency in Sichuan province by DEA-Malmquist index method. The study results indicate that the overall level of industrial green efficiency in Sichuan province is not high, but with a gradually increase trend in recent years. The reason is mainly caused by technology efficiency, but the technology progress index retardants the promote effect of technology efficiency. In the four major economic regions, the industrial green efficiency of Panxi heads the list, next are Chengdu Plain of Sichuan and South Sichuan, the worst is Northeast Sichuan.

Acknowledgements. This work was supported by the Key Program (Grant No. 16AGL003) of the National Social Science Foundation of China. And we thank the support of the Social Science Planning Annual Project of Sichuan Province (SC16C011), Soft science research plan project of Sichuan province (2017ZR0098), Soft science research project of Chengdu city (2015-RK00-00156-ZF), The Central University Basic Scientific Research Business Expenses Project of Sichuan University (skqy201623).

References

1. Chen C, Han J, Fan P (2016) Measuring the level of industrial green development and exploring its influencing factors: empirical evidence from China’s 30 provinces. Sustainability 8(2):153
2. Eiadat Y, Kelly A et al (2008) Green and competitive? an empirical test of the mediating role of environmental innovation strategy. J World Bus 43(2):131–145
3. Fare R, Grosskopf S, Norris M (1994) Productivity growth, technical progress, and efficiency change in industrialized countries: reply. Am Econ Rev 84(5):1040–1044
4. Fleiter T, Fehrenbach D et al (2012) Energy efficiency in the german pulp and paper industry-a model-based assessment of saving potentials. Energy 40(1):84–99
5. Fowlie M, Reguant M, Ryan S (2016) Market-based emissions regulation and industry dynamics. J Polit Econ 124:249–302
6. Honma S, Hu JL (2014) A panel data parametric frontier technique for measuring total-factor energy efficiency: an application to japanese regions. Energy 78:732–739
7. Jia RM (2011) The ecological character of evolution in sichuan province and regulatory strategy. Ind Technol Econ 8:98–104
8. Mei S, Wang LJ, Li WB (2009) Efficiency assessment of coal industry sustainable development by improved data envelopment analysis. In: International conference on management science and engineering, pp 1830–1835
9. Mukherjee R, Sengupta D, Sikdar SK (2015a) Sustainability in the context of process engineering. Clean Technol Environ Policy 17(4):833–840

10. Mukherjee R, Sengupta D, Sikdar SK (2015b) Sustainability in the context of process engineering. Clean Technol Environ Policy 17(4):833–840
11. Nagesha N (2008) Role of energy efficiency in sustainable development of small-scale industry clusters: an empirical study. Energy Sustain Dev 12(3):34–39
12. Ng R, Low JSC, Song B (2015) Integrating and implementing lean and green practices based on proposition of carbon-value efficiency metric. J Clean Prod 95:242–255
13. Su L, Zheng H, Wang Y (2013) Evaluation on green development of China's provincial industry. Resour Environ (8):116–122
14. Tu Z, Xiao G (2009) Research on China's industrial growth model under environmental constraints. J World Econ 11:41–54
15. Vasauskaite J, Streimikiene D (2016) Energy efficiency development in lithuanian furniture industry. Adv Manag Sci 4:9–15
16. Wang C, Huang XJ (2006) On the evaluation and diagnosis of obstacle for circular industrial economy in Jiangsu Province. Ind Eng Manag 112(2):301–315
17. Wu J (2009) Tfp growth and convergence across China's industrial economy considering environmental protection. J Quant Tech Econ 11:17–27
18. Xuxiao WU (2016) Research on the improving path and dynamic evaluation of regional industrial green development efficiency: Taking the heavy chemical industry zone of Qinghai, Henan and Fujian as examples. Ecol Econ 2:63–68
19. Yang Y (2012) Research on evaluation of sichuan low-carbon economy efficiency. China Popul Resour Environ 6:52–56
20. Yi H, Liu Y (2015) Green economy in China: regional variations and policy drivers. Glob Environ Change 31:11–19
21. Zhang C, Liu H (2014) The impact of unbalanced regional development on industrial green total factor productivity: a new perspective based on three stages dea adjustment efficiency. Contemp Econ Res 9:39–45
22. Zhang J, Zhu L (2012) Research on technological innovation efficiency of industrial enterprises based on green growth of regions in China. J Quant Tech Econ 2:113–125
23. Zhang M, Zhou ZF (2010) Research on dynamic evaluation on the independent innovation capacity of high-tech industries in China. Soft Sci 95:242–255

Risk Control

Machine Learning and Neural Network for Maintenance Management

Alfredo Arcos Jiménez(✉), Carlos Quiterio Gómez Muñoz, and Fausto Pedro García Márquez

Ingenium Research Group, Castilla-La Mancha University, Ciudad Real, Spain
Alfredo.Arcos@alu.uclm.es

Abstract. The paper presents a novel approach that allows to optimize the ultrasonic wave sensors for a condition monitoring system employing. It can detect and diagnosis different faults with a signal, such as delamination, mud or ice on blades of wind turbines. This methodology allows to avoid the redundancy of sensors, since a specific number of ultrasonic transducers can determine the structural condition using guided waves. The signal is pre-processed with the aim of removing the noise, then extracted and selected features to be later classified by Machine Learning and Neural Networks. Finally, for each damage or anomaly, the best classifier will be evaluated. The best classifier of each damage will act on a parallel network that will process the signal sent by the sensor.

Keywords: Non-destructive testing · Fault detection and diagnosis · Condition monitoring system · Wavelet transforms · Machine learning · Neuronal network

1 Introduction

The objective in the next decades is to obtain a competitive and sustainable energy to all countries, allowing to reduce dependence on fuels to households, industries and transportation. Wind power and Concentrated Solar Power (CSP) are being two of the main renewable energy sources. Its importance in the energy market is being essential, and the forecasts show it will continue in the near future.

Renewable energy industry requires significant improvements in reliability, lifetime or availability that it is done by an efficient maintenance based on Condition Monitoring Systems (CMS). CMS is the process of determining the condition of system [37,38,49], where one of the main proposes is to identify changes between two main states of the structure, damaged and undamaged [36,39,41]. They will provide different patterns that will be used in order to analyze the condition of the system [19].

A complex CMS contains numerous sensors that generates different type of signals with information and sampling frequencies [41]. It needs to be processed, preferable on-line, with a minimum computational cost and with high accuracy to reduce false alarms [39].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_115

This paper present a novel approach to reduce the number of sensors, maximizing the signal processing analysis to detect different damages. The guided wave ultrasonic signals are employed for fault detection and classification.

Structural Health Monitoring (SHM) is a technology that combines advanced CMS, together with signal processing, to determine the condition of the structures on line or not [19,36,41]. SHM leads to increase the Reliability, Availability, Maintainability and Safety (RAMS) of the system [16,41]. SHM allows also to know the different levels of the defect severity. It will be useful for an optimal maintenance management to reduce costs and increase the profitability.

Farrar and Sohn [16] considered pattern recognition in SHM. This methodology consists in four stages shown in Fig. 1.

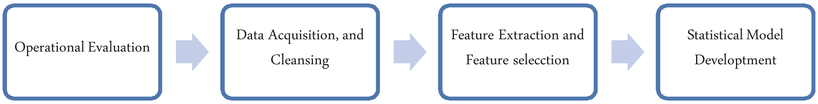


Fig. 1. Methodology SHM

The problems caused by the implementation of damage identification system is assessed in operational evaluation phase. Data acquisition and normalization involves the choice of sensors and their location, the normalization procedure of the data collected and the selection of the data to be used in the feature selection process. Dimensionality reduction techniques have been developed to remove irrelevant data (noise) and redundant features, and categorized mainly into Feature Extraction (FE) and Feature Selection (FS). FE contains information of the signal that distinguishes between a damaged or non-damaged structures [57]. FS selects the data that are better for the detection of damage and executes a condensation of the data [16,40]. Finally, classifiers, based on statistical models, calculate and cluster the dataset depending on the damage. These classifiers are grouped as: Supervised Learning (SL) and Unsupervised Learning [35].

This paper proposed patterns recognition of each damage by Machine Learning and Neural Network. Then approach minimize the number of sensors, where through an ultrasonic signal emitted by a sensor is able to classify different states of different damages or anomalies.

This process is complex because each damage must be analyzed individually. On the other hand, there is no generic classifier that provides the best results. Even the same classifier for the same signal depends on the methodology of FE and its subsequent FS.

2 Approach

Figure 2 shows the schematic approach for determining the level of damage or anomaly based on SHM employing ultrasonic guided waves.

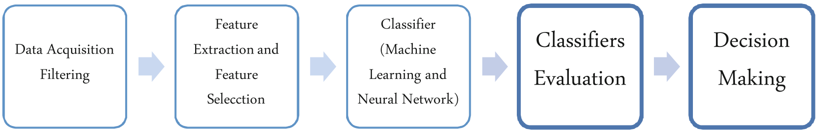


Fig. 2. Methodology schedule for determining the level of damage or anomaly

2.1 Signal Pre-Processing

Sensors generate noise from a variety of sources and therefore a signal pre-processing is necessary to eliminate/reduce the dataset that does not provide useful informatio [45]. Standard statistical techniques have been employed in references [10,53]. Yu et al. [64] were able to reduce global noise using averaging techniques and Daubechies Wavelet (DW) to eliminate local high-frequency perturbations. Denoising and compression signal in Guided Waves (GW) based on the Discrete Wavelet Transform (DWT) was employed by Rizzo and Scalea [55]. There are a large number of research and reviews articles on filtering in the treatment of ultrasonic guided waves [46]. Hamming [21] performs a review about low-pass filters available for data smoothing. This paper will consider Wavelet transforms for filtering the signals.

The denoising of the signal is performed employing a multilevel 1-D wavelet analysis using Daubechies family. The wavelet decomposition structure of the signal is extracted. The threshold for the de-noising is obtained by a wavelet coefficients selection rule using a penalization method provided by Birgé-Massart [47]. An overly aggressive filtering could eliminate data that should show, for example, small echoes that come from defects. Figure 3 shows the original signal and the de-noised signal when it is applied the wavelet de-noised filter. The Wavelet de-noising filter does not produce an unwanted signal delay in contrast to other digital filters.

It is observed that the filter removes noise significantly, and does not eliminate information that is related to different structural features.

2.2 FE Methods

FE affects the learning and classification process. FE and FS are capable of improving learning performance, presenting less computational complexity and historical dataset [58].

Classical dimensional reduction techniques such as Principal Component Analysis (PCA) [26] and Multidimensional Scale (MDS) [29] are extensively employed. However, the Linear and Nonlinear Regression systems are used when the signal type is complex and the first order interactions are not enough to derive good results [20].

Two autoregressive models have been applied in this paper for FE: Linear (AR) and nonlinear (NARX). AR is a lineal mathematical model commonly employed for FE because of its high computational performance. AR method

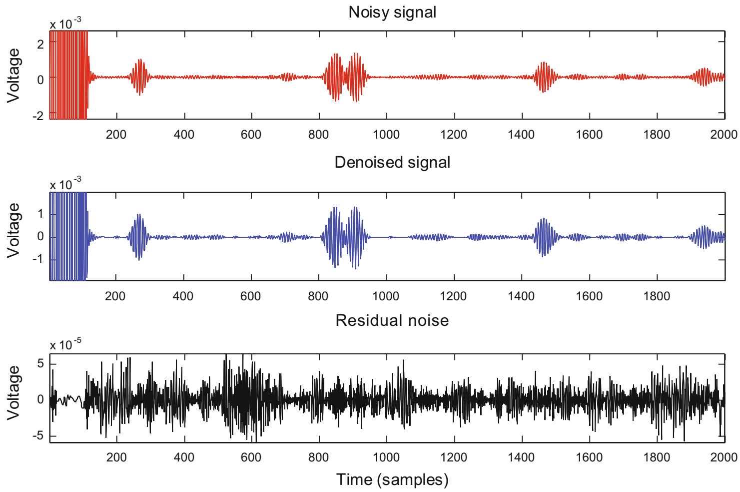


Fig. 3. Decomposition detail five (D5), De-noised D5 and extracted residual noise

is popular because estimation of AR parameters (features) is achieved easily by solving linear Yule-Walker equations [13]. NARX presents a better behavior in the most of the nonlinear systems. NARX was introduced by Leontaritis and Billings [30] and it has been widely used in the identification of nonlinear dynamic models and feature selection [6]. NARX presents a good generalization capability, as effective training algorithm, a fast convergence rate, and can represent a wide class on nonlinear systems [3,18,25].

Within the NARX model, the lineal-in-the-parameters structure is a method widely used in nonlinear dynamics models. It is a linear combination of model terms, or basis functions, that are some nonlinear-usually polynomial- functions of the system variables [32]. In this case, the monomials of the linear combination are the features. The linear-in-the-parameter models are well structured to adaptive learning, with learning conditions and demonstrable convergence and capacity parallel processing, in addition to numbers applications in signal processing and pattern recognition [24,36].

2.3 FS Methods

FS techniques can be categorized into three approaches: embedded, wrapper or filter [33]. The main methods for FS are Information Gain, Relief, Fisher Score and Lasso [27]. In this paper, the main objective to reduce dimensionality and it is not intended to select the best features.

SHM literature propose several methods to select the AR model order: AIC [2]; Akaike's Final Prediction Error [1] (FPE); Partial Autocorrelation Function (PAF) [60], and; Root Mean Square (RMS) [31]. AIC statistical technique has been used to reduce the dimensionality of the feature extraction. The AIC is one of the most efficient techniques for order optimization [35]. The AIC is a measure of the goodness-of-fit of an estimated statistical model, based on the trade-off between fitting accuracy and the number of estimated parameters.

2.4 Classifiers

(1) Machine Learning Approach

① Decision Trees

Decision Tree (DT)is a classifier used in many fields to study if the data contains different classes of objects that can be interpreted significantly in the context of a substantive theory [34,48,51,52].

DT generates a split of space from a labeled training set. The objective is to separate the elements of each class into different labeled regions (leaves) minimizing the local error. Each internal node in the tree is a question (decision) that determines which branch of the tree must be taken to reach a leaf. DT is determined for; (1) how to split the space (Splitting Rules); (2) stop condition of splitting; (3) labeling function of a region, and; (4) measurement of error.

The purpose of the Splitting Rule is to minimize the impurity of the node. The recursive splitting algorithm stops when it finds any of the following conditions: the node is pre-set maximum deep; all elements of the node are same class; there is no empty sub node or; SR does not reach a pre-set value.

To label a leaf or region once it is already considered as a terminal is considered to develop a DT. Equation (1) establishes the labeling function.

$$l' = \arg\min l' \left\{ \sum_{l=1}^{k} N_I \times c_{l,l'} \right\}, \tag{1}$$

where N_I is the number of elements of class l in the region, l' is the class to label and $c_(l, l')$ is the labeling cost.

$c_(l, l')$ with all classes is calculated and l' is selected, which minimizes the error. The label that minimizes the error of a region is the most populated class. In case of a tie, a random one is chosen. Equation (2) gives the classification average error

$$\varepsilon = \frac{1}{N} \sum_{l=1}^{N} \mathrm{err}(l, l'_i), \tag{2}$$

where $\mathrm{err}(l, l'_i)$ is the error of labeling a class l as l'. This error is solved by splitting the space and assigned to each split a label.

② Discriminant Analysis (DA)

The ultrasonic signals considered are homogeneous for the same frequency and, therefore, the classifier that provides better results is LDA [15]. This classifier is valid for ultrasonic signals obtained at different frequencies and used for speech recognition. This type of classifier is based on a geometrical approach. A linear function representing a decision limit divides a feature space into regions that have common properties. The aim of this method is to find the expression of linear discriminant functions that enable the objects classification in the considered classes.

③ Support Vector Machine (SVM)
SVM [9] is a supervised multivariate classification method. "Supervised" refers to a training step where the algorithm learns the differences between pre-specified groups to be classified [59]. SVM treats each feature as a point in a high-dimensional space, being the number of dimensions the same to the number of rating levels. Each feature is assigned to a group and the linear classification function (3) learns the characteristics to discriminate among five groups. A limit of decision, or hyperplane (a generalization of a plane of $n-1$ dimensions which divides an n-dimensional space), must be defined to separate the data based on class membership and to classify linearly the dataset. However, for a linearly separable problem, there are an infinite number hyperplanes correctly classified data. SVM algorithm finds the optimal one characterized by the largest margin between classes. The margin is defined as the distance of the closest training data points of the hyperplane. These points are the most difficult to classify and they are called support vectors. The hyperplane is defined by a weight vector, which is a linear combination of the support vectors, and specifies both a direction and a displacement which together define the maximum margin classifier. The decision function $D(x)$ is given by Eq. (3),

$$D(x) = w\emptyset(x) + b, \tag{3}$$

where w and b are the SVM parameters, and $\emptyset(x)$ is a kernel function

The hyperplane is defined by Eq. (3), and the distance between the hyperplane and pattern x can be written by Eq. (4).

$$\frac{D(x)}{\|w\|}. \tag{4}$$

A training classifier is designed to find the value of w that maximizes the margin between the class boundary and the training patterns. The objective function of the training algorithm is given by Eq. (5).

$$J = \frac{\|w\|^2}{2}. \tag{5}$$

In this case, the linear function Kernel and one vs one method have been employed [43].

④ k-Nearest Neighbors
k-Nearest Neighbors(k-NN) [11,12] is a high-performance classifier widely used in Machine Learning [42]. k-NN rule classification is an extension of the Nearest-Neighbor (NN) rule. Given a set x of the k samples, the rule assign to each sample is to label most frequently represented among the k nearest samples [15].

k-NN search technique and k-NN based algorithms are widely used as benchmark learning rules. The relative simplicity of the k-NN search technique makes it easy to compare the results from other classification techniques to k-NN results.

The accuracy of k-NN classification significantly depends on the metric used to compute distances between different samples [61]. In most cases, the best performing classifier is Fine k-NN [62], using metric Euclidean distance in Eq. (6).

$$d_{st}^2 = (x_s - y_t) \cdot (x_s - y_t)' \tag{6}$$

⑤ Ensemble Subspace Discriminant (ESD)
Ensemble methods are learning algorithms that construct a set of classifiers whose individual decisions are combined in some way (generally by weighted or unweighted voting) to classify the set of features of each pattern [14]. Bagging, boosting, and random subspaces are general techniques that can be used with any type of base classifier. Ensemble Bagged Tree (EBT) and Random Subspace Method (RSM) have been the most successful methods. EBT uses the Breiman's 'random forest' algorithm [7]. RSM is a parallel learning algorithm proposed by Ho [23].

(2) Artificial Neuronal Network (ANN)
ANN unidirectional supervised through a MLP with training by backpropagation algorithm [56] has been applied. Backpropagation with algorithm scaled conjugate gradient and performance Cross Entropy [28,44] with 'Early Stopping' to avoid overfitting [50] has been training mode. ANN is given by Eq. (7),

$$z_k = \sum_j w'_{kj} y_j - \theta'_i = \sum_j w'_{kj} f\left(\sum_i w_{ji} x_i - \theta_j\right) - \theta'_j, \tag{7}$$

where x_i is ANN input, y_i is hidden layer output, z_i is final layer output, t_k is targets output, $w_j i$ is hidden layer weight, w'_{kj} is final layer weight, θ_j is hidden layer bias, θ'_k is final layer bias and $f(.)$ is the activation function of sigmoid type, employed as the activation function of the ANN.

The MLP process tests initially with an ANN architecture and is trained with 70% of the total of the experiments and 30% of them to test the network. Then the ANN architecture is chosen according to the accuracy and performance. Finally, ANN is tested with different cases (30%) to know if the learning is right and to the accuracy.

Backpropagation (BP) is one of the simplest and most general methods for supervised training of multi-layer ANNs. Scaled, standardization, normalization that perform pre-processing inputs techniques have been employed to accelerate the BP training. Scaled conjugate gradient and performance Cross Entropy are employed to increase the computational learning cost of ANN.

2.5 Evaluation Classifier

Analysis Receiver Operating Characteristic (ROC), as Confusion Matrix (CM), is employed to evaluate of the classification. CM determines the accuracy of a classifier and measures its performance. The main parameters in CM are:

True Positive (TP) : corresponds to real successes of the classifier
False Positives(FP) : is the sum of the values of a class in the corresponding CM column, excluding the TP
False Negative(FN) : is the sum of the values of a class in the corresponding CM row, excluding the TP
True Negatives(TN) : is the sum of all columns and rows, excluding that class's column and row

Recall function, R, given by Eq. (8), provides the probability to be correctly classified, called true positive rate or hit rate:

$$R = \frac{TP}{TP + FN}. \tag{8}$$

Specificity, S, also called the true negative rate, measures the proportion of negatives that are correctly identified as negatives.

$$S = \frac{TN}{FP + FN}. \tag{9}$$

Precision, P:

$$P = \frac{TP}{FP + TP}. \tag{10}$$

F-score, F [54]

$$F = \frac{2 \times P \times R}{P + R}. \tag{11}$$

There are two conventional methods to establish the average performance in all categories: macro-averaging and micro-averaging [63]:

Macro-average: P^M, R^M and F^M is obtained are the mean of all P_i^M, R_i^M and F_i^M where M denotes Macro-average, and i is the scenario, then for each category they are calculated, i.e. they are valued are evaluated locally P_i^M, R_i^M and F_i^M, and then globally P^M, R^M and F^M.

Micro-average: P^M, R^M and F^M value is obtained as: i) TP_i, FP_i, FN_i values is calculated for each of the scenario; ii) the value of TP, FP and FN are calculated as the sum of TP_i, FP_i and FN_i, and; iii) to applying the equation of the measure that correspond.

There are several indices extracted from the ROC curve to evaluate the efficiency of a classifier. Area Under Curve (AUC) represents the area between the ROC curve and the negative diagonal [5,22], being between 0.5 and 1. Values ≤ 0.5 indicates that the classifier is invalid and a value of 1 indicates a perfect rating because there is a region in which, for any point cut, the values of R and Pare equal to unit. The statistical property of AUC is equivalent to the Wilcoxon test of ranks [22]. The AUC is also related to the Gini coefficient [8], which is twice the area between the diagonal and the ROC curve.

The recommendations by Demsar [4], and the extensions by Garcia and Herrera [17], have been employed to perform the comparative analysis of classifiers. Friedman Test will be used to test the null hypothesis that all classifiers achieve the same average. Bonferroni-Dunn Test is applied to determine significant differences between the top-ranked classifier and the following. Holm Test is used to contrast the results.

3 Approach Scheme

The methodological process is represented schematically in Fig. 4. Firstly, for each damage or anomaly, the best classifier is selected.

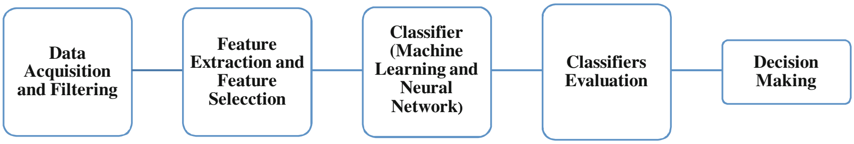


Fig. 4. Methodology flow for determining the classifier

The methodological process will follow the scheme given in Fig. 5 when the best classifier for each damage is set.

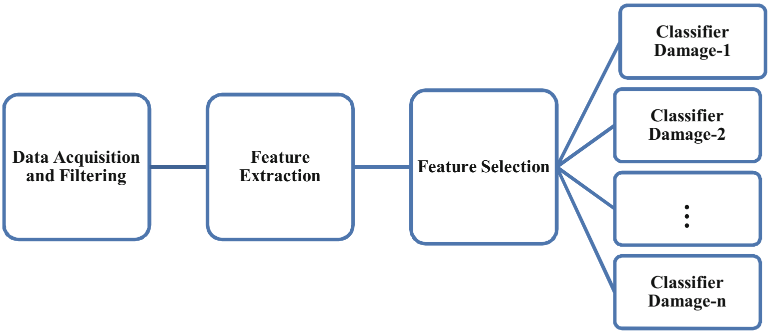


Fig. 5. Methodology flow for determining the damage

4 Conclusions

The paper presents a novel approach to optimize the sensors in a condition monitoring system employing ultrasonic waves. The approach can detect different potential faults with a single signal emitted by a sensor, such as delamination, mud or ice on blades of wind turbines. This methodology allows to avoid the redundancy of sensors, since a specific number of ultrasonic transducers can determine the structural condition using guided waves. The signal is pre-processed with the aim of removing the noise, then extracted and selected features to be then classified by Machine Learning and Neural Networks. Finally, for each damage or anomaly, the best classifier will be evaluated. The best classifier of each damage will work in a parallel network that will process the signal sent by the sensor.

Acknowledgements. The work reported herewith has been financially supported by the Spanish Ministerio de Economía y Competitividad, under Research Grant DPI2015-67264-P and RTC-2016-5694-3.

References

1. Akaike H (1969) Fitting autoregressive models for prediction. Ann Inst Stat Math 21(1):243–247
2. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723
3. Alazrai R, Lee CSG (2012) An narx-based approach for human emotion identification. In: Ieee/rsj international conference on intelligent robots and systems, pp 4571–4576
4. Ar J (2006) Statistical comparisons of classifiers over multiple data sets. J Mach Learn Res 7(1):1–30
5. Bradley AP (1997) The use of the area under the roc curve in the evaluation of machine learning algorithms. Pattern Recognit 30(7):1145–1159
6. Brankovic A, Falsone A et al (2016) Randomised algorithm for feature selection and classification. Xiv preprint
7. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
8. Breiman L, Friedman J, et al (1984) Classification and regression trees. CRC Press
9. Burges CJC (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc 2(2):121–167
10. Chopra I (2002) Review of state of art of smart structures and integrated systems. AIAA J 40(11):2145–2187
11. Cover T, Hart P (1967) Nearest neighbor pattern classification. IEEE Trans Inf Theory 13(1):21–27
12. Dasarathy BV (1990) Nearest neighbor (nn) norms: nn pattern classification techniques. Los Alamitos IEEE Comput Soc Press 13(100):21–27
13. De Lautour OR, Omenzetter P (2010) Damage classification and estimation in experimental structures using time series analysis and pattern recognition. Mech Syst Signal Process 24(5):1556–1569
14. Dietterich TG (2000) Ensemble methods in machine learning. In: International workshop on multiple classifier systems, pp 1–15
15. Duda RO, Hart PE, Stork DG (2012) Pattern classification. Wiley
16. Farrar CR, Doebling SW, Nix DA (2001) Vibration based structural damage identification. Philos Trans R Soc B Biol Sci 359(359):131–149
17. Garc S, Herrera F (2008) An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. J Mach Learn Res 9(12):2677–2694
18. Ghosh S, Maka S (2009) A narx modeling-based approach for evaluation of insulin sensitivity. Biomed Signal Process Control 4(1):49–56

19. González-Carrato RRDLH, Márquez FPG, Dimlaye V (2015) Maintenance management of wind turbines structures via mfcs and wavelet transforms. Renew Sustain Energy Rev 48:472–482
20. Guyon I, Elisseeff A (2006) An introduction to feature extraction. Springer, Heidelberg
21. Hamming RW (1989) Digital filters, 3rd edn. Prentice Hall
22. Hanley JA, Mcneil BJ (1982) The meaning and use of the area under a receiver operating characteristic (roc) curve. Radiol 143(1):29–36
23. Ho TK (1998) The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 20(8):832–844
24. Hong X, Mitchell R et al (2008) Model selection approaches for non-linear system identification: a review. Int J Syst Sci 39:925–946
25. Jiang C, Song F (2010) Forecasting chaotic time series of exchange rate based on nonlinear autoregressive model. In: Advanced Computer Control (ICACC), 2010 2nd international conference, pp 238–241
26. Jolliffe I (2002) Principal component analysis. Wiley Online Library
27. Jovic A, Brkic K, Bogunovic N (2015) A review of feature selection methods with applications. In: International convention on information and communication technology, electronics and microelectronics, pp 1200–1205
28. Kroese DP, Rubinstein RY et al (2013) Cross-entropy method. In: Encyclopedia of operations research and management science. Springer, pp 326–333
29. Kruskal JB, Wish M (1978) Multidimensional scaling. Book on Demand Pod
30. Leontaritis IJ, Billings SA (1985) Input-output parametric models for non-linear systems part ii: stochastic non-linear systems. Int J Control 41(2):303–328
31. Levinson N (1949) The wiener (root mean square) error criterion in filter design and prediction. Stud Appl Math 25(1–4):261–278
32. Li K, Peng JX, Bai EW (2006) A two-stage algorithm for identification of nonlinear dynamic systems. Automatica 42(7):1189–1197
33. Liu H, Motoda H (2008) Less is more. Computational methods of feature selection. In: H Liu, e Motoda, H (eds) Chapman and Hall/CRC, pp 3–17
34. Loh WY, Shih YS (1997) Split selection methods for classification trees. Stat Sinica 7(4):815–840
35. Lordo RA (2001) Learning from data: Concepts, theory, and methods. Technometrics 43(1):105–106
36. Márquez FPG, Munoz JMC (2012) A pattern recognition and data analysis method for maintenance management. Int J Syst Sci 43(6):1014–1028
37. Márquez FPG, Pedregal DJ, Roberts C (2013) New methods for the condition monitoring of level crossings. Int J Syst Sci:878–884
38. Márquez FPG, Pardo IPG, Nieto MRM (2015) Competitiveness based on logistic management: a real case study. Ann Oper Res 233(1):157–169
39. Márquez FPG, Pérez JMP et al (2016) Identification of critical components of wind turbines using fta over the time. Renew Energy 87:869–883
40. Martinez-Luengo M, Kolios A, Wang L (2016) Structural health monitoring of offshore wind turbines: a review through the statistical pattern recognition paradigm. Renew Sustain Energy Rev 64:91–105
41. Marugán AP, Márquez FPG, Pérez JMP (2016) Optimal maintenance management of offshore wind farms. Energies 9(1):46
42. Michalski RS, Carbonell JG, Mitchell TM (2013) Machine learning: an artificial intelligence approach. Springer Science and Business Media

43. Milgram J, Cheriet M, Sabourin R (2006) "one against one" or "one against all": which one is better for handwriting recognition with svms? In: Proceedings of international workshop on frontiers in handwriting recognition
44. Møller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw 6(4):525–533
45. Muñoz JMC, Márquez FPG, Papaelias M (2013) Railroad inspection based on acfm employing a non-uniform b-spline approach. Mech Syst Signal Process 40(2):605–617
46. Munoz CG, Márquez FG (2016) A new fault location approach for acoustic emission techniques in wind turbines. Energies 9(1):40
47. Munoz CQG, Márquez FPG, Tomás JMS (2016) Ice detection using thermal infrared radiometry on wind turbine blades. Measurement 93:157–163
48. Pal M, Mather PM (2003) An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sens Environ 86(4):554–565
49. Papaelias M, Cheng L et al (2016) Inspection and structural health monitoring techniques for concentrated solar power plants. Renew Energy 85:1178–1191
50. Prechelt L (1998) Automatic early stopping using cross validation: quantifying the criteria. Neural Netw Off J Int Neural Netw Soc 11(4):761
51. Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106
52. Quinlan JR (2014) C4. 5: programs for machine learning. Elsevier
53. Raghavan AC, Cesnik CES (2007) Review of guided-wave structural health monitoring. Shock Vib Dig 39(2):91–114
54. Rijsbergen CJV (1979) Information retrieval. Butterworth-Heinemann
55. Rizzo P, Scalea FLD (2004) Discrete wavelet transform to improve guided-wave-based health monitoring of tendons and cables. Proc SPIE Int Soc Optical Eng 5391:523–532
56. Rumelhart DE, McClelland JL et al (1988) Parallel distributed processing, vol 1. IEEE
57. Staszewski WJ, Robertson AN (2007) Time-frequency and time-scale analyses for structural health monitoring. Philos Trans R Soc A Math Phys Eng Sci 365(1851):449
58. Tang J, Alelyani S, Liu H (2014) Feature selection for classification: a review. In: Documentación Administrativa, pp 313–334
59. Vapnik VN (1999) An overview of statistical learning theory. IEEE Trans Neural Netw 10(10):988–999
60. Wei WWS (1994) Time series analysis. Addison-Wesley publ Reading
61. Weinberger KQ, Saul LK (2009) Distance metric learning for large margin nearest neighbor classification. J Mach Learn Res 10(1):207–244
62. Xu Y, Zhu Q et al (2013) Coarse to fine k nearest neighbor classifier. Pattern Recognit Lett 34(9):980–986
63. Yang Y (1999) An evaluation of statistical approaches to text categorization. Inf Retr J 1(1):69–90
64. Yu L, Bao J, Giurgiutiu V (2004) Signal processing techniques for damage detection with piezoelectric wafer active sensors and embedded ultrasonic structural radar. Proc SPIE Int Soc Optical Eng 5391:492–503

Volatility Spillover Between Foreign Exchange Market and Stock Market in Bangladesh

Shibli Rubayat[1](✉) and Mohammad Tareq[2]

[1] Faculty of Business Studies, University of Dhaka, Dhaka 1000, Bangladesh
shibli@du.ac.bd

[2] Department of Accounting and Information Systems, University of Dhaka, Dhaka 1000, Bangladesh

Abstract. This paper investigates the link between foreign exchange market and stock market in Bangladesh through volatility spillover between the markets. It also examines the volatility persistence and asymmetric effect of information on the volatility of these two financial markets. Using data from January 1, 2009 to December 12, 2016 Taka/USD exchange rate's volatility and the stock return's volatility of CSE General Index this paper utilizes ARCH and GARCH models to investigate the spillover, persistence and asymmetry effect of volatility. The results reveal high level of presence of volatility of CSE General Index but not of the Taka/USD exchange rate. The result also finds asymmetric effect of information i.e. bad news is followed by higher volatility compared to that of after good news in the stock market. However, the result shows that volatility in the CSE General Index spillover that of the Taka/USD exchange rate but not vice versa. The meaning and significance of the study is also presented in the paper.

Keywords: Foreign exchange market · Stock market · Volatility · Spillover · Taka/USD (The local and United States Currency)

1 Background of the Study

Foreign exchange market and capital market are linked in many countries and show inter-dependence [1,7,11]. Among others, volatility of these markets links each other and spillover from one market to another. The volatility is a key factor that causes risk in these markets and thus influences economic activities in a country. Volatility information of these markets is used for risk management, portfolio allocation and stock trading strategies [12].

Bangladesh, a south Asian country struggling for poverty reduction and economic development, has a smaller and less equipped capital market than its neighboring countries [10]. A developed capital market can aid a country's economic development by turning savings into investment. Researchers have consistently found a positive relationship between a developed stock market with economic growth and capital accumulation [8,9]. As mentioned above, knowledge

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_116

about the volatility help investors in their risk management, portfolio allocation and stock trading strategies and is considered one of the ways of developing a capital market. However, there is no notable study in Bangladesh that has investigated the volatility of these financial markets, especially how these two markets are linked through volatility.

The contribution of the research lies in the fact that no notable study has been done so far that investigates the link between the volatility of the capital market and foreign exchange markets in Bangladesh. It is hoped that this study will inform policy-making decisions in relating to these two market and help investors in making investment decision.

1.1 Methodology

The study uses CSE General Index of Chittagong Stock Exchange (the 2^{nd} largest in the port city) from January 1, 2009 to December 1, 2016. The official exchange rate of Taka/USD of the same period was also used. These data were collected from Chittagong Stock Exchange and from the central bank of Bangladesh, 'Bangladesh Bank' respectively. The Taka/USD exchange rates and the CSE General Index for the sample period are shown in Figs. 1 and 2 respectively.

The return of the stock market was calculated as follows:

$$R_t = \ln\left(P_t/P_{t-1}\right) \times 100.$$

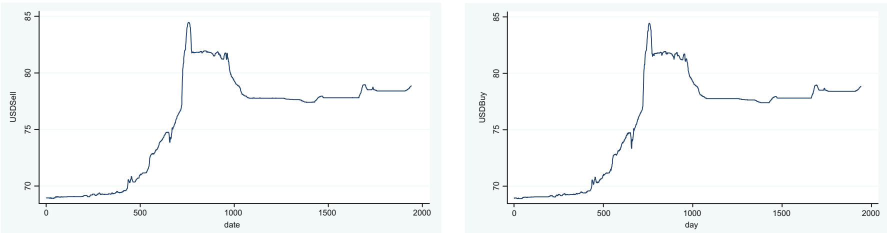


Fig. 1. Taka/USD exchange rate (From January 1 2009 to December 12, 2016)

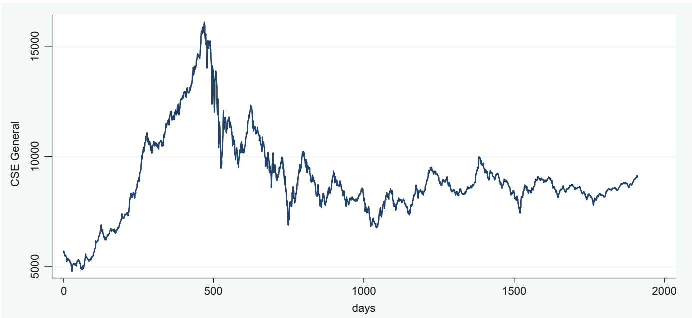


Fig. 2. CSE General (From January 1, 2009 to December 12, 2016)

Daily spread between buy and sell rate of Taka/USD was used to measure return on Taka/USD exchange rate. The Taka/USD exchange rates return and the CSE General Index daily return is for the sample period are shown in Figs. 3 and 4 respectively.

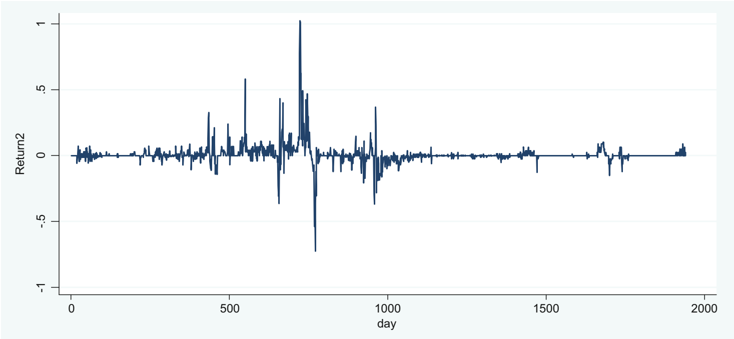


Fig. 3. USD/Taka exchange rate return (Dec 12, 16 to Jan 1 2009)

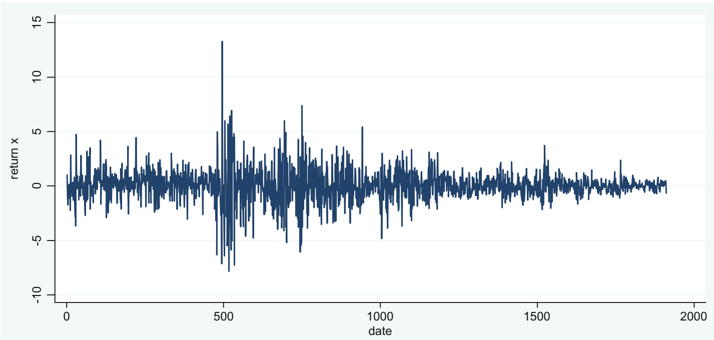


Fig. 4. CSE General Index stock return (Dec 12, 16 to Jan 1 2009)

To analyze volatility, asymmetric volatility and spillover effect, different variation of the Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) were used in this study. Maximum Likelihood estimation method was used to estimate the parameters of the GARCH models of this study. Augmented Dickey–Fuller (ADF) unit root tests for the CSE General Index return and Taka/USD exchange rate return series were done and the results reject the hypothesis of unit root. GARCH is a model for time-series where volatility is clustered together. The autoregressive term in a GARCH model takes into account past values on the time-series. Heteroskedasticity means that data contains non-constant variance, which by running a standard OLS would result in erroneous standard errors and confidence intervals. GARCH models corrects these treat heteroskedasticity as a variance inside the model.

The ARCH model was the forerunner to the GARCH specifications. Before this, volatility used to be measured using rolling standard deviations over past time periods. However, the problem with this approach is that each time period receives equal weight in the estimation of the current period's variance. The ARCH process, proposed by Engle [5], solves this by letting these weights be parameters to be estimated. This allows the data to determine the best weight to be used on the present volatility [6]. An ARCH model is presented below:

$$r = \pi + \varepsilon_t,$$
$$\sigma_t^2 = c + \sum_i^q \alpha_i \varepsilon_{t-i}^2,$$

where $\varepsilon_t = \sigma_t z_t$; z_t is an i.i.d random variable with mean zero and variance one and σ^2 is the squared conditional variance in the model.

So the conditional variance is here described as a distributed lag of past squared innovation [3]. That means that the variance of the current error term is a function of the size of the previous period's squared error term. In order to avoid a very large number of coefficient in high order polynomial, Bollerslev [2] developed the GARCH-model as a generalization of Engle's [5] ARCH-model. This makes a declining weight that never reaches zero [6]. GARCH $(1,1)$ is the most common model to use in empirical research. A GARCH $(1,1)$ model is presented below:

$$r = \pi + \varepsilon_t,$$
$$\sigma_t^2 = c + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2,$$

where, $\varepsilon_t = \sigma_t z_t$; z_t is an i.i.d random variable with mean zero.

2 Persistence and Asymmetry of Volatility

The study utilize a GARCH $(1,1)$ model for the CSE General Index return and Taka/USD exchange rate return. The result shows that the "$\alpha + \beta$" estimated from the CSE General Index is about .99. This means that volatility is highly persistent for the CSE General Index. It means that if there is volatility in the CSE stock market it decays slowly over a long period of time. However the study did not find the persistence of volatility in the Taka/USD rate (See Table 1).

In the world of financial market the reality is that negative news often has much bigger impact on volatility than good news has on relaxing these fluctuations i.e. negative shocks at time $t-1$ have a stronger impact on the variance at time t than positive shocks. The GARCH model is symmetrical, negative and positive shocks are given equal weight. The study also investigated the asymmetry effect of positive news and negative news on the market volatility using EGARCH-model. Result of the asymmetric GARCH shows that bad news causes

Table 1. Persistence and asymmetry of CSE General Index and Taka/USD exchange rate

	CSE General Index	CSE General Index	Taka/USD exchange rate
Con	.005*(.027)	.005*(.027)	0.00**(00)
α	.16***(.014)	.16***(.014)	1.23(1.15)
β	.83***(.011)	.83***(.011)	.34***(.03)
γ (asymmetric term)	-	$-$.02*(.011)	$-$.6(1.16)
n	1911	1911	1940

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

higher volatility ($\gamma = -0.02^{*}$).compared to good news in the stock markets in Bangladesh (See Table 1).

Past studies have shown that volatility in one market can affect other markets [4,11]. The study also investigates whether the volatility crated in stock market spills over the other and vice versa. To study this, MGARCH is used. The result shows that volatility in the CSE General Index spillover that of the USD/Taka exchange rate but not vice versa. Table 2 shows the spillover effect of CSE General Index on the Taka/USD exchange rate market.

Table 2. Spillover effect of stock market on foreign exchange market

	CSE General Index	Taka/USD exchange rate
Volatility Taka/USD	$-$.02(.40)	-
Volatility CSE General Index	-	.0007***(.0002)
N	1911	1940

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

3 Conclusion

This study investigates and analyses the link between the stock market and foreign currency (ForEx) market of Bangladesh. The study was done to get insight about on the volatility and its spillover effect between the stock and the foreign exchange markets, and consequently the degree of their integration to expand the information set available to international portfolio managers, multinational corporations and policymakers for decision-making and policy formulation.

It was found that the persistence of volatility in stock market in Bangladesh is very high which indicates that volatility clusters in stock market and decays slowly. This study is similar to studies to Wu [11] and Jebran and Iqbal [7]. However, the study found low persistence in Taka/USD currency market. The study also found that there is asymmetry in the volatility of in the stock market i.e. negative news causes higher volatility compared to good news in the market. The result shows that volatility in the CSE General Index spillover that of the Taka/USD exchange rate but not the opposite.

References

1. Bala DA, Asemota JO (2013) Exchange-rates volatility in Nigeria: application of garch models with exogenous break. J Appl Stat 4(1):89–116
2. Bollerslev T (1986) Generalised autoregressive conditional heteroskedasticity. Econometrica 31(302):2–27
3. Campbell JY, Lo AW et al (1997) The econometrics of financial market. Princeton University Press, Princeton
4. Cheung YW (2001) Equity price dynamics before and after the introduction of the euro: a note. Multinational Financ J 5(2):113–128
5. Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50(4):987–1007
6. Engle RF, Patton AJ (2001) What good is a volatility model? Quant Financ 1(2):237–245
7. Jebran K, Iqbal A (2016) Dynamics of volatility spillover between stock market and foreign exchange market: evidence from Asian countries. Financ Innov 2(1):1–20
8. Levine R (1999) Financial development and economic growth: views and agenda. J Econ Lit 35(2):688–726
9. Levine R, Zervos S (1996) Stock markets, banks, and economic growth. Am Econ Rev 88(3):537–558
10. Siddiqui J (2010) Development of corporate governance regulations: the case of an emerging economy. J Bus Ethics 91(2):253–274
11. Wu RS (2005) International transmission effect of volatility between the financial markets during the Asian Financial Crisis. Transition Stud Rev 12(1):19–35
12. Zivot E (2009) Practical issues in the analysis of univariate GARCH models. Springer, Heidelberg

Cost/Efficiency Assessment of Alternative Maintenance Management Policies

Diego Ruiz-Hernández(✉) and Jesús María Pinar-Pérez

Department of Quantitative Methods, CUNEF-Ingenium, Madrid, Spain
d.ruiz@cunef.edu

Abstract. In this work we compare a number of maintenance scheduling policies for a collection of deteriorating machines. We consider a framework where a manager must allocate the effort of a small number of technicians for providing preventive maintenance to a larger number of machines. The aim is mitigating wear and preventing failures or breakdowns, guaranteeing the system's availability, reliability and safety. We evaluate different maintenance regimes that emerge in practice, distinguishing two main groups: time based and condition based policies. The comparisons are conducted by means of an intensive computational assessment of the performance of the different policies discussed.

Keywords: Machine maintenance management · Condition based maintenance · Time based maintenance · Index policies · Simulation

1 Introduction

Maintenance interventions in deteriorating equipment are aimed at guaranteeing its availability, reliability and safe performance whilst maximizing output and minimizing waste. Such interventions may be predictive, preventive, or corrective. Regarding the planning horizon, whereas corrective maintenance is reactive (interventions come as a response to a failure), predictive and preventative interventions are typically scheduled using either time-based or condition-based schemes. Time-based interventions are usually scheduled according to an age-based regime or following a pre-determined calendar (usually as indicated by the producer or by certain legal requirements). In this work we focus our attention on preventative maintenance interventions.

There is a vast amount of literature addressing the problem of scheduling preventative maintenance interventions in one single machine. Some recent references are [3–6, 12, 13]. Good reviews of earlier contributions please see [1, 8, 20]. However, it is hard to find references in literature addressing the problem of scheduling maintenance interventions to more than one machine, among them [10, 15, 21]. Moreover, in most of the work addressing single machine maintenance, the attention is focused on finding optimal intervention times given the information gathered about the current condition of the machine without taking

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_117

into consideration the costs associated with the machine operation, intervention and potential breakdowns. These values become particularly relevant when it comes to deal with the design of multiple machines maintenance schedules, as the trade-off intervention/operation-breakdown becomes more relevant.

Notwithstanding the shortage of academic work, technical literature shows that maintenance tasks are typically allocated following either calendar based programmes or, alternatively, regimes based on information about the current state of the equipment [2,7,11,14,18,22]. Based in this observation, in this work we compare the performance of a number of maintenance scheduling policies which broadly replicate the approaches taking in real life problems.

2 Problem Formulation

In this work we address the problem of scheduling preventive maintenance interventions in a collection of independent machines by a limited number of technicians. The aim is operating the system at the least operation/maintenance cost whilst minimizing the cost of failures or breakdowns. At each period the manager selects the number of machines that will be maintained during that period. If necessary, technicians are allowed to remain idle during certain period.

As discussed in [19], the manager's problem can be modeled a discrete-time infinite-horizon Markov Decision Problem, where each machine is modeled as a two action (operation/intervention) Markov Decision Chain that evolves over a discrete state space[1]. Roughly speaking, during the operation phase the machine deteriorates, operation costs increase and the probability of a failure or breakdown becomes larger. Maintenance interventions, which are conducted at a cost, aim at returning the machine to some pristine or as-good-as-new state. The objective of the manager is to minimize the discounted operation/intervention cost of equipment over an infinite horizon.

In order to characterize the decision maker's problem we introduce the following notation: let β, M and R represent the discount factor, and the number of machines and repairmen, respectively. The system's state space is represented by the set $\mathscr{S}$, with the state at any decision epoch t given by $\mathbf{S}(t) = \{s_m(t), m = 1, \cdots, M\} \in \mathscr{S}$. The actions available are represented by the set $\mathscr{A}$, whose elements $\mathbf{a} = \{a_m : a_m \in 0, 1, m = 1, \cdots, M\}$ represent the admissible actions at any given state $\mathbf{S}(t)$; where $a_m = 1$ indicates that machine m must be intervened during the current decision period, and $a_m = 0$ prescribes operation. Finally, $\mathscr{C}$ is the collection of all possible vectors $\mathbf{C}^{\mathbf{a}}(\mathbf{S})$, representing the costs incurred from taking action $\mathbf{a} \in \mathscr{A}$ when the system is in state $\mathbf{S} \in \mathscr{S}$.

With this elements we can now write down the objective of the decision maker as the one of choosing the stationary, deterministic policy σ that minimizes

[1] Please notice that, even if the machines' condition or level of wear may not necessarily be a discrete variable, if the deterioration of the machine occurs smoothly over a linear space, then the state space can be discretized or uniformised in order to obtain a discrete state space [17].

the expected discounted operation/maintenance cost incurred over an infinite horizon, namely $V^* = \min_\sigma\{E_\sigma[\sum_{t=0}^{\infty} \beta^t \mathbf{C}^{\mathbf{a}(t)}\mathbf{S}(t)]\}$.

Unfortunately, this problem suffers a dimensionality problem making standard dynamic programming techniques unavailable for finding an optimal solution. Instead, in real life, different approaches are taken (see [2,14,22]) in order to obtain the most efficient maintenance scheduling regime possible. Most of these approaches fall within the five broad categories described in the following section.

3 Maintenance Policies

In this section we briefly describe the five maintenance regimes that have been used in our work. All of them aim at minimizing the overall operation/maintenance cost of equipment. Three of them are condition based and presume the existence of a condition monitoring system that provides timely information about the system's state. Most of the condition based policies that are used in real life fit (with minimal adaptations) within the more general frameworks described below. However, because of the large investment costs involved, in some cases condition monitoring systems may not economically feasible (see, for example, [16]). In such cases time based regimes are particularly useful. For this reason, the fourth policy discussed in this work is a purely calendar based one. Finally, the last policy discussed below combines features of both, condition based and time based regimes.

(1) Threshold Policy

At each decision epoch, the manager observes the state of the system and assigns maintenance tasks to those machines whose state is equal to or above than certain threshold. If the number of candidate machines is larger than the number of technicians, then the effort is allocated to those machines with larger state of wear. Ties are broken randomly. More precisely, the action set of the threshold policy for a threshold T at time t is given by $\mathbf{a}(t) = \{a_m(t) : s_m(t) \leq T, \sum_{m=1}^{M} a_m(t) \leq R\}$.

(2) Pure Condition Based Policy

At each decision epoch, this policy prescribes to intervene those machines with larger observed state of wear. As in the previous case, ties are broken randomly. This kind of policies do not take into account other characteristics of the machines as age, deterioration rate, operation and intervention costs, etc. If we define an ordering $s_{m,(i)}$, $i = 1, \cdots, M$, where $s_{m,1}$ is the machine with largest state and so on, then we can build the following two sets $\bar{a} = \{s_{m,(i)}, i = 1 \cdots, R\}$ and $\underline{a} = \{s_{m,(i)}, i = R, \cdots, M\}$, therefore the condition based policy is represented by: $\mathbf{a}(t) = \{a_m | a_m = 1 \,\forall\, m : s_{m,(i)} \in \bar{a};\; a_m = 0 \;\forall\, m : s_{m,(i)} \in \underline{a}\}$.

(3) Index Policies

Index based families of maintenance scheduling policies have been proposed in order to address the deficiencies of purely condition based policies, [9,10]. These policies exploit all the information available for each machine at any time in order to compute a so-called activity or intervention index. The indices are calibrated for each machine/state pair and the policy prescribes to intervene those machines with larger indices. Typically, the indices will have negative values for early states of wear, when the machine is supposed to be in an as-good-as-new condition. Consequently, only machines with positive indices will be intervened. This result has been extended to all the other policies evaluated in this article, imposing that no machine must be intervened whenever it lies in a state whose index has negative value[2].

If we let Im, s_m represent the index of machine m at state s_m, and $\mathscr{I}$ is the R^{th} largest index, then the actions available at time t are given by $\mathbf{a}(t) = \{a_m | a_m = 1 \quad \forall\, m : Im, s_m > \mathscr{I}; \quad a_m = 0 \quad \forall\, m : Im, s_m \leq \mathscr{I}\}$.

(4) Calendar Based Policy

Periodical or calendar based policies follow a pre-defined intervention plan. In our setting, a periodical intervention time is fixed for all the machines, with the first intervention scheduled after certain sojourn time (when the machine is operated from new). Even though in practice this calendar is determined by the acquisition date and the periodical interventions prescribed by the maker, in our formulation the interventions calendar is determined as follows: Once the sojourn period is finished, at each period the decision-maker intervenes as many machines as technicians are available according to their level of wear (pristine machines are not intervened). Once a machine has been intervened, a new intervention is scheduled at certain time in future (the interval between interventions is the same for all machines and exogenously determined). Once all the machines have been scheduled and intervened, a new cycle begins. For the sake of results's comparability with the index and other pure condition-based regimes, a variant of this policy consisting of preventing the machine to be intervened in early states of wear, has been included. These states coincide with the minimal state for which all the indices computed for the index policy are non-negative.

(5) Dynamic Calendar Policy

When the collection of machines includes equipment from different versions, ages or makers, their deterioration rates may differ and a rigid calendar-based policies may not be completely appropriate. In order to address these cases, we propose an extension of the pure calendar-based policy which is an hybrid of calendar and condition based policies: at each decision epoch, the less deteriorated machines in the plan are substituted for other -more deteriorated ones- which are not in the current period's plan. The machines whose intervention is postponed join a queue or *buffer* of delayed machines that will be included in future plans.

This policy, depicted in Fig. 1, works as follows. First, all the machines are assigned an intervention period as in the *Calendar Based Policy.* Once the intervention stage has started, at each period the manager checks on the condition of

[2] Please notice that this condition has only been imposed for the sake of results' comparability and does not apply for more general cases.

all the machines. If any of the machines scheduled for intervention at the current period is still at the pristine state, the machine is dropped-off from the list and its identifier is included in the buffer. Simultaneously, if there are non-pristine machines in the buffer, they are immediately included, by order of wear, in one of the empty slots of the current plan. The decision maker then checks on the condition of all the other, non-scheduled, machines and identifies those units whose state of wear is worst than the best machine in the plan. If no machines are identified, the intervention stage starts and the maintenance tasks are deployed. If, otherwise, some machines are picked during the check, they are either allocated to any empty slot in the current plan (if any), or substitute any scheduled machines which are in a better condition. The machine (or machines if there are ties) with the largest state of wear in the current intervention plan is never substituted. The displaced machines are included in the buffer and the intervention actions are deployed. As it was done for the purely calendar based policy, a variant of this policy consisting of preventing the machine to be intervened in early states of wear, has also been considered.

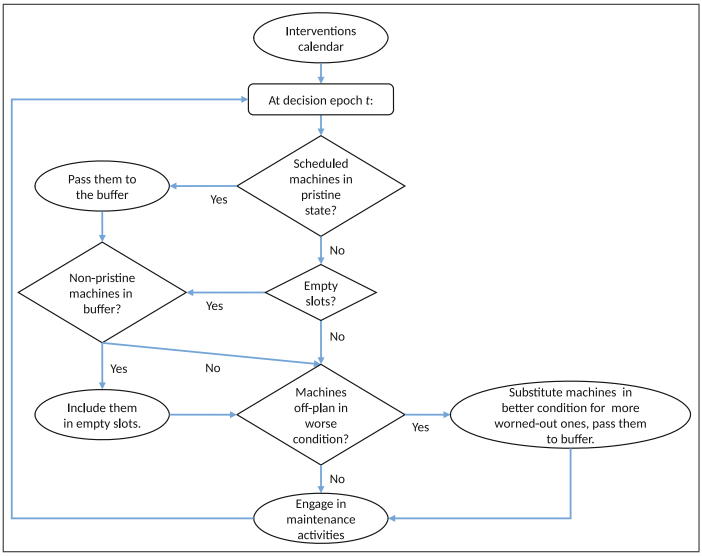


Fig. 1. Dynamic calendar policy

4 Numerical Experiments

A number of numerical experiments was conducted in order to compare the performance of each of the policies discussed in Sect. 3 under a collection of different scenarios characterized by the following variables: maintenance and operation costs; breakdown costs; transition rate; breakdown probability; and number of machines technicians.

The value of the objective function, V^* is computed by means of the simulation of 1040 decision periods. If we consider that interventions are conducted weekly, this accounts for a planning horizon of 20 years, which corresponds to the lifespan of a typical industrial equipment under normal conditions.

A total of 30 instances were simulated for each of 19440 possible scenarios. The scenarios were obtaining as combinations of the following parameter specifications. Instances consisted of 30, 60, 100 and 150 machines with the number of technicians given by ceiling of the 2%, 4% and 6% of the number of machines. The machines evolve over a 52 states deterioration space, visiting a new state every 4 to 6.67 weeks in one case, and every 2 to 4 weeks in another. Therefore, the deterioration probabilities are random extractions from $U(0.15, 0.25)$ and $U(0.25, 0.50)$ distributions, respectively. Additionally, the deterioration probabilities increase 0.5% in every transition. For the breakdown probabilities we considered five cases where each machine is first assigned a random value B_m obtained from one of the following distributions $U(0.05, 0.10,), U(0.10, 0.15), U(0.15, 0.20), U(0.20, 0.25)$ and $U(0.25, 0.30)$[3]; once this value is fixed, the breakdown probability of machine m at state s_m is computed as $B_m e^{0.025 s_m}$.

Regarding the costs, three possible values were considered for the breakdown costs: 500, 1500 and 3000. A random linear operation costs function was defined with three alternative intersects obtained from distributions $U(45, 55), U(135, 165)$ and $U(270, 330)$, and slope extracted from a $U(11.25, 18.75)$ distribution. Finally, twelve different configurations of the intervention costs where defined: three of them, constant random functions with support on $U(90, 110), U(225, 275)$ and $U(450, 550)$; and other nine quadratic configurations with the intersect following the same distribution as the constant cases and the quadratic term obtained from one of the following distributions: $U(0.0075, 0.0125), U(0.0225, 0.0375)$ and $U(0.0375, 0.0625)$.

In order to determine the states where intervention is allowed for all the policies, a parameter *Forbid* was introduced. This parameter takes the value of the minimal state for which all the indices in the index policy take non-negative values for each particular instance. In order to do so, the indices corresponding to the index policy are computed in the startup routine of each instance of the numerical experiments. These values were obtained based on [19] by means of the expression $I(m, s_m) = G(s)(1 - E[\beta^{\tau(0,s)}]) - K(0, \tau(0, s)) - E[\beta^{\tau(0,s)} C(\tilde{s})]$. In these expressions, with $G(s) = \frac{K(s,1)+E[\beta C(s^*)]-C(s)}{1-\beta}$. In these expressions, $E[\beta^{\tau(0,s)}]$ represents the accumulated discount during a transition from state 0 to state s, which takes an expected time given by $\tau(0, s)$; $K(0, \tau(0, s))$ represents the expected discounted intervention cost incurred in a transition between states 0 and s; $E[\beta^{\tau(0,s)} C(\tilde{s})]$ is the expected discounted cost of intervening a machine after $\tau(0, s)$ periods of operation ($\tilde{s}$ represents the expected arrival state after the operation sojourn); $K(s, 1)$ is the expected discounted cost of operating during one period when the initial state of the machine is s; $E[\beta C(s^*)]$ represents the expected discounted cost of intervening the machine at the expected arrival state

[3] Notice that in each case all the random values are extracted from the same distribution.

Table 1. Results of the illustrative example

Criteria	Index policy	Condit. Bsd policy	Calendar policy	Calendar-F policy	Dyn.Cal. policy	Dyn. Cal.-F policy	Threshold policy
Average St.	2.25	2.26	2.93	3.33	3.30	3.37	2.26
Var. Av. St.	7.91e−02	1.24e−02	2.52e−01	2.05e−01	3.95e−02	4.35e−2	1.24e−2
Maximum St.	8	8	18	20	10	16	8
Breakdowns	2889	2878	2934	2956	2621	2671	2818
Av. St. Interv.	5.90	5.91	5.77	7.85	8.61	8.86	5.91
Var. Av. Interv.	3.69e−01	1.10e−02	8.24e−01	1.06e+00	1.61e−01	1.96e−01	1.08e−02
Total Cost	14220580	14293900	15610959	16366256	16140478	16310495	14293900
Numb. Interv.	3118	3118	3120	3120	3120	3120	3118

† The benchmark example consists of a collection of 60 machines and three technicians. The operation costs are linear with constant parameter 150; the intervention costs are quadratic, with intersect 500 and quadratic coefficient 0.03. The breakdown costs are 1500, with breakdown parameter 0.10. Parameter *Forbid* takes value 3.

after one period operation (notice that $s^* \in (0, s, s+1)$); finally, $C(s)$ is the cost of intervention when the machine is in s.[4]

Once the indices are obtained, the threshold policy is computed fixing the threshold equal to the value *Forbid*; likewise, the pure condition based policy is only allowed to intervene machines with current state larger than *Forbid*. For the calendar based policies two alternatives approaches were taken: in the first one, the procedure was allowed to intervene any machine with state different than zero; in the second one, intervention was only admitted for states larger than *Forbid*.

A summary of the results obtained in a benchmark example is shown in Table 1. The columns represent the seven alternative policies, with the suffix $(-F)$ indicating that parameter *Forbid* has been used. The columns show, respectively, the average visited state and its variance, the maximum visited state, the number of breakdowns over 1040 periods, the average intervened state and it variance, the total operation/maintenance cost and the total number of interventions. The cost structure and the indices are illustrated in Fig. 2.

Figure 3 illustrates the evolution of an insulated machine under two of the most relevant policies: Index and Calendar. It can be seen that, under the index policy, the interventions tend to occur, in average, at state number 5, whereas the evolution under the calendar policy tends to be more erratic, as this policy does not take into consideration the state of the machine for scheduling interventions.

Given that the numerical values of the results of the experiments are not as important as the patterns detected during the numerical computations, we limit ourselves to presenting a graphical summary of our main findings. Firstly, we notice that the average cost of the operation/maintenance schedule over the

4 Closed form expressions for this equations, which have been omitted for the sake of briefness, are available from the authors upon request.

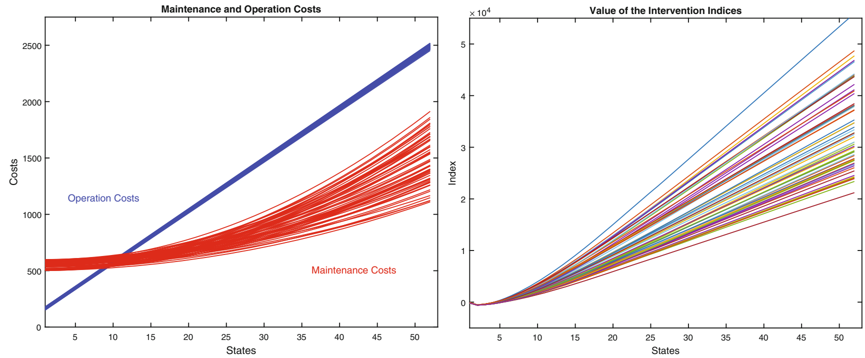


Fig. 2. Costs and Indices for 56 states in the illustrative example

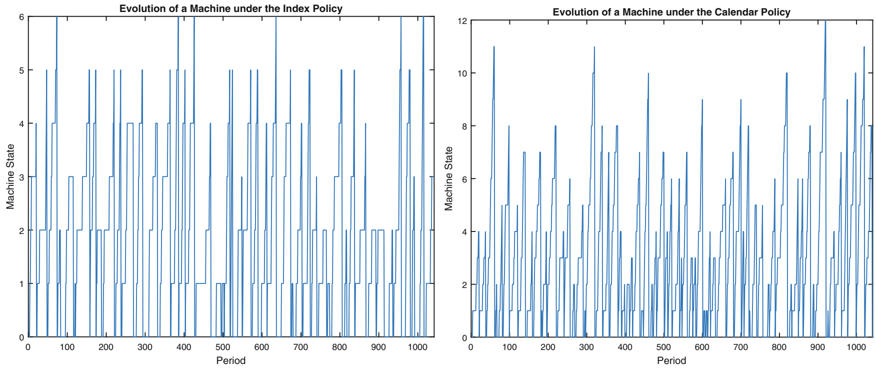


Fig. 3. Evolution of a random machine under index and calendar policies.

20 years planning horizon follows a very similar pattern for all policies, with the time dependent policies showing larger costs (in average) than the state dependent ones. This can be seen in panel (a) of Fig. 4, where the index (blue), Calendar-F (red) and the threshold policies (green) are compared[5]. This panel shows the total cost incurred by the seven policies over the 19440 scenarios and (the values are averages over 30 repetitions for each case)[6].

The sketch suggest the existence of four well defined blocks, corresponding to each of the four different specifications for the number of machines. Each of these main blocks is composed by three more or less differentiated segments, which

[5] For the sake of simplicity, the dynamic calendar policies and the simple version of the calendar policy have not been illustrated in this figure. However, their behavior does not present significant differences with respect to the results shown in the displays.

[6] Each point in the graph corresponds to the average cost, computed over 30 randomly generated problems, of a particular combination of operation, maintenance and breakdown costs, deterioration rates, breakdown probabilities and number of machines and repairmen.

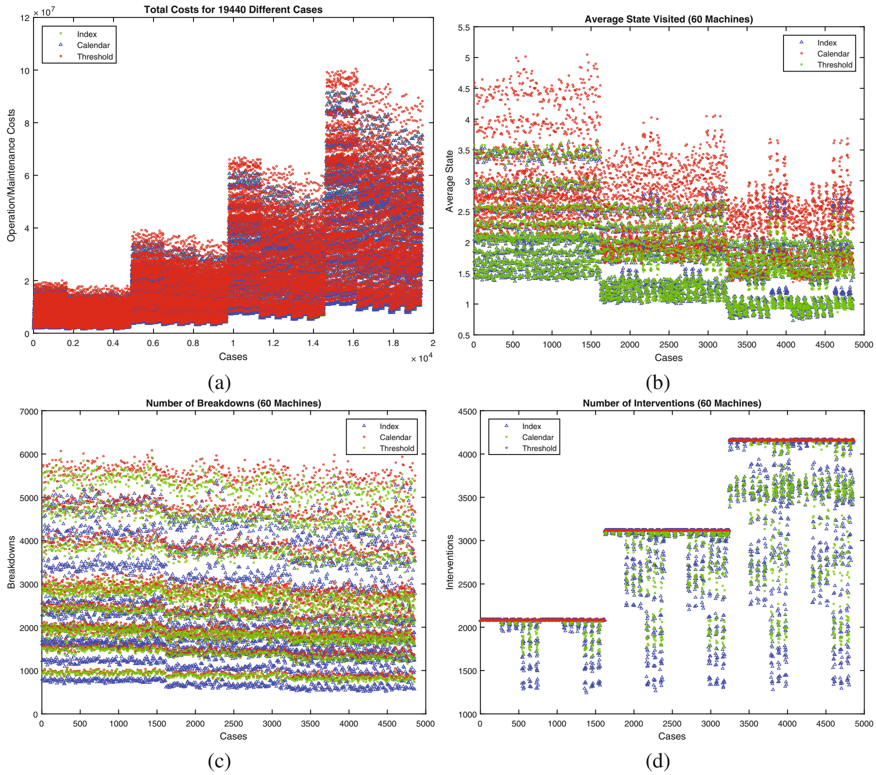
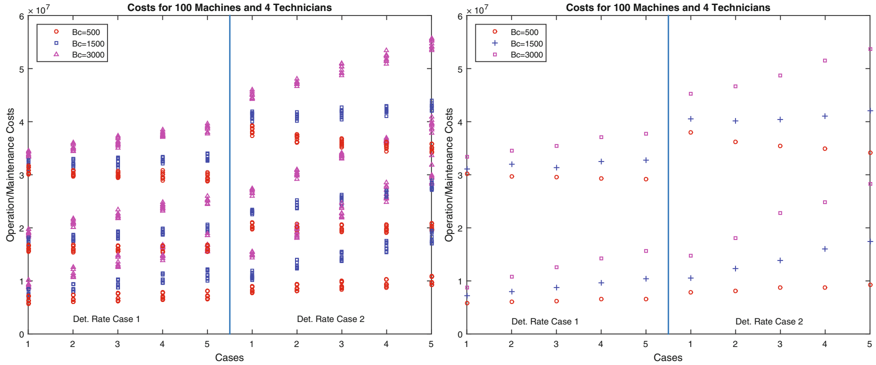


Fig. 4. Costs performance of the index, calendar-f and threshold policies over a 20 years planning horizon

correspond to the three possible number of technicians used in our experiments. The figure clearly shows that the costs are increasing on the number of machines and decreasing on the number of technicians. In order to make our discussion simpler, and given the fact that the results observed in one group follow the same structure as the others (differing mainly in their absolute scale), in panels (b), (c) and (d) we limit our discussion to the results observed for the cases with 60 machines. Panel (b) shows the average visited state during the planning horizon. It can be appreciated that the time based regimes tend to visit larger states (i.e. machines deteriorate more) than state based policies. It can also be observed that the behavior of the threshold and the index policy is quite similar. This suggests that, even in the absence of a well grounded index heuristic, a decision maker will-in general-do better with a threshold policy than with a time based one. In panel (c) we show the number of breakdowns suffered by the system over the planning horizon. Again, the calendar policy shows a larger number of breakdowns, and the index policy shows a smaller number of failures than the other two. Finally, panel (d) depicts the number of interventions conducted by each policy over the 20 years horizon. It can be appreciated that condition based policies intervene much less than the time based ones. Moreover, the index policy clearly outperforms the threshold one. It is also interesting to see that the calendar policy intervenes a constant number of machines irrespectively of the specific cost structure of the particular scenario, with interventions depending only on the number of machines and technicians available.



(a) (b)

Fig. 5. Inner structure of the index policy results

We finally take a look inside of our plots in order to illustrate the inner structure of the results. In order to do so, we pick the case where there are 100 machines, four technicians and the costs are quadratic. Figure 5 illustrates the corresponding 810 results obtained for the index policy. Vertically, each panel is divided in two sections, corresponding to each of the two alternative deterioration rates proposed. Each section consists of five columns representing each of the possible configurations of the breakdown probabilities. Horizontally, it is possible to distinguish three blocks consisting of three groups of points (red, blue and magenta). Each of these groups is associated with a different value of the breakdown cost as illustrated by the legend. Additionally, each of the three main levels (consisting of three colors each) correspond to a different operation cost. For example, the southernmost red group corresponds to the smallest operation and breakdown costs; the next one (blue) corresponds to the lowest operation and the second breakdown cost, and so on. It can be seen that the colored points are grouped in small clusters. Each of these clusters correspond to the nine different combinations of the maintenance cost parameters. The second panel depicts the cases of a fixed maintenance cost and two alternative (high and low) operation costs, together with the three alternatives for the breakdown cost. This panel allows a more clear visualization of the behavior of our formulation. It can be seen, for example, that a larger deterioration rate implies larger costs for all the configurations (comparing the right and the left hand side sections). In general, the costs are larger as the likelihood of a breakdown increases, however, it is worth noticing that for low breakdown costs, this relationship may become negative. This may be reflecting an incentive to intervene less when the replacement cost of the equipment is small compared with the maintenance costs. This structure can be observed all over the set of experiments with only small variations between policies.

5 Conclusion

In this article presented a collection of general purpose machine maintenance scheduling policies that mimic the intervention strategies that are used in practice. These policies can be grouped in two broad categories: time and condition based policies. We additionally suggest a novel alternative consisting of a mixture of both. Our aim has been to obtain an assessment of their relative performance before different combinations of operation, intervention and failure costs, deterioration rates and breakdown probabilities.

With this aim, we conducted a large number of experiments. In general the results are according to what intuition will suggest: condition based policies in general perform better than time based ones. We also found that threshold policies in general outperform the pure condition based ones and are almost as efficient as the more sophisticated index policies. Our results show that when intervention and breakdown costs are low, time based policies may perform better than the index one, as this will try to postpone the intervention tasks, increasing the probabilities of breakdowns.

Our results are consistent with the trend that has been observed in the last few years in heavy equipment based industries: the use of time based policies has been gradually relegated in favor of more efficient condition based strategies. However, time based policies are still relevant because of both guarantee, safety and maintainability purposes, and the fact that in many cases condition monitoring systems require an important investment that cannot always be afforded by small firms.

References

1. Ahmad R, Kamaruddin S (2012) An overview of time-based and condition-based maintenance in industrial application. Comput Ind Eng 63(1):135–149
2. Ben-Daya M, Ait-Kadi D et al. (2009) Handbook of maintenance management and engineering, vol 7. Springer
3. Dao CD, Zuo MJ, Pandey M (2014) Selective maintenance for multi-state series-parallel systems under economic dependence. Reliab Eng Syst Saf 121:240–249
4. Do P, Voisin A et al (2015) A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions. Reliab Eng Syst Saf 133:22–32
5. Doostparast M, Kolahan F, Doostparast M (2014) A reliability-based approach to optimize preventive maintenance scheduling for coherent systems. Reliab Eng Syst Saf 126:98–106
6. Fitouhi MC, Nourelfath M (2014) Integrating noncyclical preventive maintenance scheduling and production planning for multi-state systems. Reliab Eng Syst Saf 121:175–186

7. Márquez FPG, Tobias AM et al (2012) Condition monitoring of wind turbines: techniques and methods. Renew Energy 46:169–178
8. Garg A, Deshmukh S (2006) Maintenance management: literature review and directions. J Qual Maint Eng 12(3):205–238
9. Glazebrook K, Mitchell H, Ansell P (2005) Index policies for the maintenance of a collection of machines by a set of repairmen. Eur J Oper Res 165:267–284
10. Glazebrook K, Ruiz-Hernandez D, Kirkbride C (2006) Some indexable families of restless bandit problems. Adv Appl Probab 38:643–672
11. Hodge V, O'Keefe S et al (2015) Wireless sensor networks for condition monitoring in the railway industry: a survey. IEEE Trans Intell Transp Syst 16(3):1088–1106
12. Khatab A, Aghezzaf EH (2016) Selective maintenance optimization when quality of imperfect maintenance actions are stochastic. Reliab Eng Syst Saf 150:182–189
13. Liu B, Xu Z et al (2014) A value-based preventive maintenance policy for multi-component system with continuously degrading components. Reliab Eng Syst Saf 132:83–89
14. Mobley R, Higgins L, Wikoff D (2008) Maintenance engineering handbook. McGraw-Hill, New York
15. Moghaddam KS (2013) Multi-objective preventive maintenance and replacement scheduling in a manufacturing system using goal programming. Int J Prod Econ 146(2):704–716
16. Pinar-Pérez J, García-Márquez F, Ruiz-Hernández D (2016) Economic viability analysis for icing blades detection in wind turbines. J Clean Prod 135:1150–1160
17. Puterman ML (1994) Markov decision processes: discrete and stochastic dynamic programming. Wiley, New York
18. Riera-Guasp M, Antonino-Daviu J, Capolino G (2015) Advances in electrical machine, power electronic, and drive condition monitoring and fault detection: state of the art. IEEE Trans Industr Electron 62(3):1746–1759
19. Ruiz-Hernández D (2007) Indexable restless bandits. Index policies for some stochastic scheduling and dynamic allocation problems. Verlag Dr. Müller, Saarbrücken, Germany
20. Sharma A, Yadava G, Deshmukh S (2011) A literature review and future perspectives on maintenance optimization. J Qual Maint Eng 17(1):5–25
21. Tao X, Xia T, Xi L (2014) Opportunistic preventive maintenance scheduling based on theory of constraints. In: Proceedings of the IIE annual conference. Institute of Industrial Engineers-Publisher, p 230
22. Wang H (2002) A survey of maintenance policies of deteriorating systems. Eur J Oper Res 139(3):469–489

Heijunka Operation Management of Agri-Products Manufacturing by Yield Improvement and Cropping Policy

Ritsuko Aoki[1(✉)] and Hiroshi Katayama[2]

[1] Department of Industrial and Management Systems Engineering, Graduate School of Creative Science and Engineering, Waseda University, Tokyo, Japan
may-mar.1103@ruri.waseda.jp

[2] Faculty of Science and Engineering, Waseda University, Tokyo, Japan

Abstract. The environment of agricultural businesses in Japan is changing drastically due to decreasing self-sufficiency rate and the number of farmers, increasing average age of concerned workers, and competition with imported products. To cope with these trends, national government revised agricultural land law which came into effect in 2009 and some companies in this industry are implementing industrial management methods such as lean management for better performance. Besides the general common problems of manufacturing industry, agri-products have distinctive features such as production in natural environment, perishability of products, demand volatility although products are living necessaries. Based on the above cognition, this study discusses implementation of Heijunka (the method of level production and supply) operation, which is one important concept of lean management.

Keywords: Lean management · Supply chain · Agri-businesses · Heijunka production · Seeding schedule · Segmented farmland

1 Introduction

Because of the decreasing self-sufficiency rate and the number of farmers, increasing of the average age of farmers and limited agricultural land, the environment of agricultural business is in tough phase [6]. In addition, extreme narrowness of agricultural land area hinders production cost reduction, which leads to decline in price competitiveness of domestic agricultural products against imported goods. For this reason, in recent years the government has considered the ways of agricultural reform and promotion of more efficient production activities. As the result, the revised agricultural land law was implemented in 2009. So many companies try to adopt the management approach developed in manufacturing industries such as TPS (Toyota Production System) and aim for stable and efficient supply. However this scientific approach against hidden problem was

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_118

not satisfactory performed in the conventional agriculture, because of managers' leadership and employees' understanding, intuitions and skills [4].

Based on the above cognition, this study investigates and analyzes cases of farmers' efforts which incorporate industrial management methods and guide to effective management of agricultural production.

2 Related Study

Vassian [7] logically derived the periodic production ordering method that achieves the minimization of inventory fluctuation by using the demand forecast values, past production ordering and the term end stock amount. Based on this, Chiyoma et al. [1] proposed the production planning model for agricultural supply chain by the following mathematical Eqs. (1), (2), and (3), where consideration is given to varied harvest quantity that peculiarly happens on agricultural products.

$$P_t = \sum_{i=1}^{L} \hat{D}_{t,t+i} - \sum_{i=1}^{L-1} P_{t-i} - I_t + SS, \tag{1}$$

$$I_t = I_{t-1} + Q_t - D_t, \tag{2}$$

$$Q_t = P_{t-L} + \varepsilon_t. \tag{3}$$

Symbols

P_t : Production quantity planned at period t, completed at period $t + L$;
$\hat{D}_{t,t+i}$: Demand quantity of period $t + i$ predicted at the end of period t;
I_t : Inventory quantity at the end of period t;
SS : Safety stock level;
L : Production lead time $+ 1$;
D_t : Demand quantity at period t;
Q_t : The yield quantity at period t, where it does not always coincide with the planned quantity;
ε_t : The difference between Q_t and P_{t-L}.

3 Background and the Production Process of the Objective Products

In this study, research procedure includes case investigation of a company that aggressively adopts lean management, model building of manufacturing process focusing on loss structure, development of Heijunka contrivance for the objective system, of which idea is cropping scheduling based on segmented farmland. Actually, this study focuses on the baby leaf growing and supply process at the collaborating company H [2]. Among the people, who carefully consider their health, baby leaf are now booming foods and this business has been growing rapidly in Japan and many SMEs including this company were established. As a result, the competitiveness of this field becomes serious. On the other hand, there is a burden on production and delivery activities because of the lack of reliable workers and useful tools/weapons regarding agricultural management.

Companies trying to mitigate such obstacles have begun to learn from companies in other industrial field and introduce their management methods and techniques. For example, TPS is a typical lean management scheme born in Japan. Company H started the TPS learning project in 2006 with Toyota Motor headquarters under supervision of former Chairperson Mr. Cho and Senior Technical Executive Mr. Hayashi, who proposed how to introduce TPS during 2006–2008 [2]. This study focuses on reaching the next stage of lean, namely realizing level production and supply (that is called Heijunka) operations based on the historical approaches such as 5S, visualization, yield improvement etc.

The objective process consists of 4 major operations, i.e. seeding, germination, harvesting and shipping, and it is illustrated in Fig. 1. Note that there are some unknown mechanism creating losses in each sub-process, these losses are called germination yield rate, harvest rate and sorting rate where weight gain is increasing quantity of the baby leaf during the growing process. Models for parameter estimation of these losses and weight gain are given in the following 4 steps with Eqs. (4)–(10) and the result is given in Table 1 [5], where it is known that germination, harvest and sorting rates are stochastically fluctuated and weight gain rate stays constant which depends on climate (warmer or cooler terms).

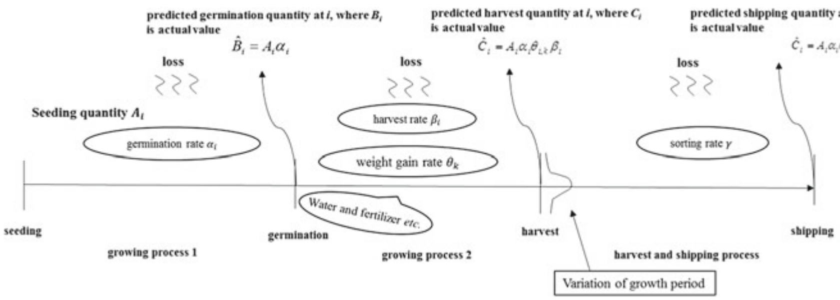


Fig. 1. Outline of the objective production process [5]

Step 1. Calculation formulas of $\hat{\delta}_i, \hat{\theta}_k$

Estimates $\hat{\delta}_i$ and $\hat{\theta}_k$ can be evaluated by least square method through minimizing the total difference between actual data of harvest and its estimated quantity as shown in Eqs. (4), (5), (6) and (7).

$$\min \sum_{i \in N} \left(\hat{C}_i - C_i\right)^2, \tag{4}$$

$$\hat{C}_i = A_i \times \hat{\delta}_i \times \hat{\theta}_k \, (k = 1, 2), \tag{5}$$

$$\text{s.t. } 0.64 \leq \hat{\delta}_i \leq 1, \tag{6}$$

$$50 \leq \hat{\theta}_k \leq 300. \tag{7}$$

Step 2. Calculation formulas of standard deviation

Estimating standard deviation of $\hat{\delta}_i$ is also important as this value represents the degree of internal turbulence that harms stable production.

$$\sigma_{\hat{\delta}} = S(\hat{\delta}_i), \tag{8}$$

where, sample data of $\hat{\delta}_i$ is obtained by the formula $\hat{\delta}_i = \frac{C_i}{A_i \hat{\theta}_k}$, which is deduced from formula (5) and S is the operator for standard deviation.

Step 3. Calculation formula of $\hat{\gamma}$

Point estimate of sorting rate $\hat{\gamma}$ can be obtained by the quotient of actual data of shipment and harvest quantities as shown in Eq. (9), where, E is the operator for mean value.

$$E(\hat{\gamma}) = E\left(\frac{D_i}{C_i}\right). \tag{9}$$

Step 4. Calculation formula of standard deviation

As the same meaning described in Step 3, standard deviation of $\hat{\gamma}$ is also important to estimate.

$$\sigma_{\hat{\gamma}} = S\left(\frac{D_i}{C_i}\right). \tag{10}$$

Symbols

A_i : Seeding quantity at planning term i [kg/a];
C_i : Harvest quantity at planning term i [kg/a];
$\hat{C}_i$: Estimated harvest at planning term i [kg/a];
D_i : Shipping quantity at planning term i [kg/a];
$\hat{D}_i$: Estimated shipping quantity at planning term i [kg/a];
N : Planning horizon [day];
α_i : Germination yield rate at planning term i;
β_i : Harvest yield rate at planning term i;
$\hat{\gamma}$: Estimated sorting yield rate;
$\hat{\theta}_k$: Estimated weight gain rate in k, ($k = 1$: warmer season, $k = 2$: cooler season)
$\hat{\delta}_i$: $\alpha_i \times \beta_i$ which is called estimated value of aggregated yield.

4 Concept of Heijunka and Its Realization

In this section, general approach and agri-specific production model for Heijunka are discussed.

Table 1. Estimated statistics of parameters [5]

	Germination rate × harvest rate	Weight gain (warmer term)	Weight gain (cooler term)	Sorting rate
Average	0.63	78.29	128.58	0.57
Standard deviation	0.14	N/A	N/A	0.21

4.1 General Approach for Heijunka

It is necessary that standard deviation and average of variable must be controlled in this order to realize continuous and stable production, which is called Heijunka production. General approach for Heijunka is illustrated in Fig. 2.

- Phase 1: Recognition of "Muri"; Firstly, current Muri must be recognized.
- Phase 2: Reduction of "Mura"; Secondly, variance reduction must be focused.
- Phase 3: Reduction of "Muda"; Finally, reduction of "Muda" must be conducted after reduction of "Mura", where level production and supply operations become much easier than before because internal turbulence was controlled in Phase 2.

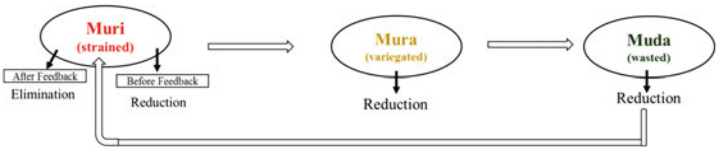


Fig. 2. General approach for Heijunka

4.2 Agri-Specific Production Model for Heijunka

For Heijunka of the objective operation, 3M ("Muri", "Mura", "Muda") are visualized and zero-tized, and seeding schedule based on segmented farmland are discussed respectively.
(1) Visualizing and zero-tizing the phenomena of 3M (Conventional lean approach)

Heijunka procedure regarding to Muri, Mura and Muda for the objective process is illustrated in Fig. 3 where target variable and its factors are shown. Stochastic simulation is a possible way to analyze the characteristics because variables and parameters in this system are fluctuated. In this study, firstly, current Muri is understood and assumed. In the second stage, variance control must be focused, which consists of 2 experiments. Factors affecting variation of harvest interval are considered as seeding interval and growth term, and that of

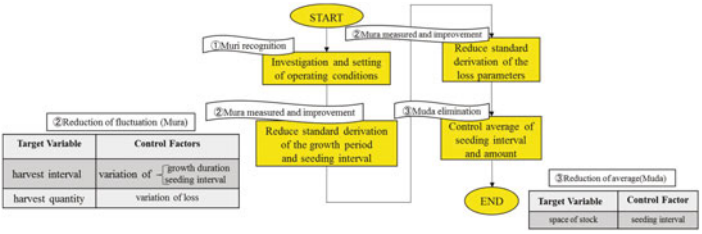


Fig. 3. Procedure of experiment

harvest quantity is considered as aggregated yield, and these standard deviation must be reduced. In the final stage, average of seeding interval is controlled under the idea of frequent small-lot production of industrial management method.
(2) Seeding schedule based on segmented farmland (Lean approach with case depending procedure)

It is essential to introduce a production system that can meet demand with limited field area because productivity improvement is indispensable in agricultural business in Japan. Therefore, seeding schedule on segmented farmland that can flexibly respond to demand quantity is considered. Design of seeding quantity, harvest quantity and harvest date is defined by Eqs. (11), (12) and (13).

$$A_i = T/N, \tag{11}$$

$$\hat{C}_i = A_i \hat{\delta}_i \theta, \tag{12}$$

$$\hat{d}_i = R_i + s_i. \tag{13}$$

Symbols
N : Number of segments of farmland;
T : Seeding capacity of entire farmland [kg];
A_i : Seeding quantity per segment at planning term i [kg/seg].
$\hat{C}_i$: Estimated harvest quantity at planning term i [kg/seg];
$\hat{\delta}_i$: Estimated value of aggregated yield of seeding quantity at planning term i;
θ : Supposed weight gain rate (percentage of weight increase);
$\hat{d}_i$: Estimated harvest date of seeded plants at term i;
R_i : Growth duration of seeds at planning term i;
S_i : Seeding date at planning term i.

In addition, an example of seeding schedule is shown in Fig. 4, and an example of input and output dates are shown in Fig. 5, where the number of segments is 9.

5 Result

In this Sect. 2 results derived from the model/procedure described in the previous chapter are summarized.

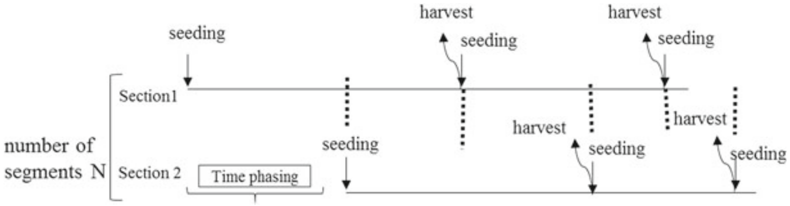


Fig. 4. Example of seeding schedule

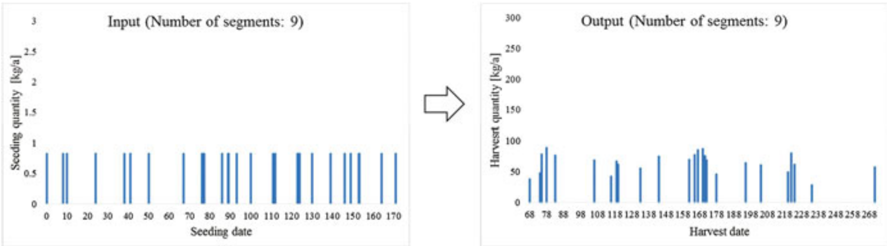


Fig. 5. Example of input and output time series

5.1 Performance of Conventional Lean Approach

(1) Phase 1: Quantification of "Muri" and its effects

The lack of human resources and skills throughout the whole process from seeding to shipment is recognized. Performance of current system is given in Table 2.

Table 2. Condition and performance of current system

	Variable			Aggregated yield	Result	
	Seeding quantity [kg/a]	Seeding interval [days]	Growth duration [days]		Harvest quantity [kg/a]	Harvest interval [days]
Average	0.85	7.12	63.26	0.63	71.27	8.28
Standard deviation	0.08	5.24	18.10	0.14	18.72	7.71

(2) Phase 2: Reduction of "Mura"

The purpose of the experiment is to evaluate the stability of harvest performance by improving fluctuation of variables. Firstly, standard deviation of the seeding interval and growth duration were reduced. Result is given in Table 3.

Secondly, in addition to the condition of the first experiment, standard deviation of the aggregated yield is reduced. Result is given in Table 4.

Table 3. Result of "Mura" reduction: first experiment (* is focused variable and performance)

	Variable			Aggregated yield	Result	
	Seeding quantity [kg/a]	Seeding interval [days]	Growth duration [days]		Harvest quantity [kg/a]	Harvest interval [days]
Average	0.89	7.05	63.26	0.63	73.35	7.06
Standard deviation	0	1.26^*	6.08^*	0.14	17.20	4.89^*

Table 4. Result of "Mura" reduction: second experiment (* is focused variable and performance)

	Variable			Aggregated yield	Result	
	Seeding quantity [kg/a]	Seeding interval [days]	Growth duration [days]		Harvest quantity [kg/a]	Harvest interval [days]
Average	0.89	7.05	63.26	0.63	72.27	7.06
Standard deviation	0	1.26	6.08	0.05^*	5.66^*	4.89

Table 5. Result of "Muda"reduction (* is focused variable and performance)

	Variable			Aggregated yield	Result	
	Seeding quantity [kg/a]	Seeding interval [days]	Growth duration [days]		Harvest quantity [kg/a]	Harvest interval [days]
Average	0.73^*	5.78^*	63.26	0.63	59.29^*	5.82^*
Standard deviation	0	1.26	6.07	0.05	4.89	4.32

(3) Phase 3: Reduction of "Muda"

The purpose of this experiment is to evaluate the influence of frequent small-lot production (Heijunka). For a pragmatic improvement point of view, a case that average seeding interval and quantity are 6/7 respectively is considered for analysis. Result is given in Table 5. Where, it is assumed that necessary farmland can be procured without any restriction but based on the contracts.

Changes in each value of standard deviation and average are described below.

- Before and after ratio of standard deviation of harvest interval is 0.56 approximately (from 7.71 to 4.32) and that of harvest quantity is 0.26 approximately (from 18.72 to 4.89).
- Before and after ratio of average of harvest interval is 0.7 approximately (from 8.28 to 5.82) and that of harvest quantity is 0.8 approximately (from 71.27 to 59.29).

Relationship of before and after improvement is given in Fig. 6. By the above procedure, harvest date and quantity can be predicted more accurately by reducing Mura which is general turbulent factor, and stock volume and space are expected to be improved by reducing both seeding interval and quantity as mentioned in Fig. 3. This procedure leads us to the Heijunka world. Worker's skill and the number of necessary workers can be encouraged to improve and then Muri of current system must be disappeared.

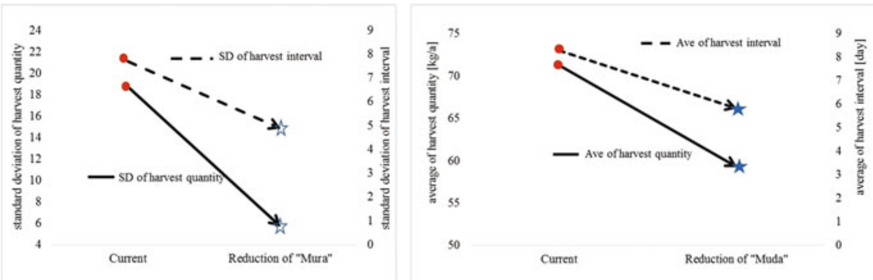


Fig. 6. Relationship on before and after improvement

5.2 Performance of Lean Approach with Case Depending Procedure

In Sect. 5.1, condition of farmland is not considered. In this section, performance of the proposed model is as described in Sect. 4.2, where farmland segmentation and its scheduling procedure [3] are added to the conventional lean approach, is evaluated by experiments. Relationship between number of segments and harvest interval and that of harvest quantity are given in Fig. 7. The star marks (★ and ☆) in these figures show the result of the model described in Sect. 5.1, where ★ and ☆ represent the performance described in Tables 3 and 4 respectively. Obtained performance of the proposed model indicates that average harvest interval and quantity are both decreasing along with increase of number of segments. Standard deviation of these variables has the same characteristics. And there are significant gaps in the characteristics between standard deviations of harvest interval obtained by Sects. 5.1 and 5.2. Namely, the proposed model discussed in this section seems to provide worse performance than conventional lean approach. Main reason will be consideration of the pragmatic restriction on segment rotation in the proposed model, which simply describes seeding is the precedence operation of harvesting in the same segment.

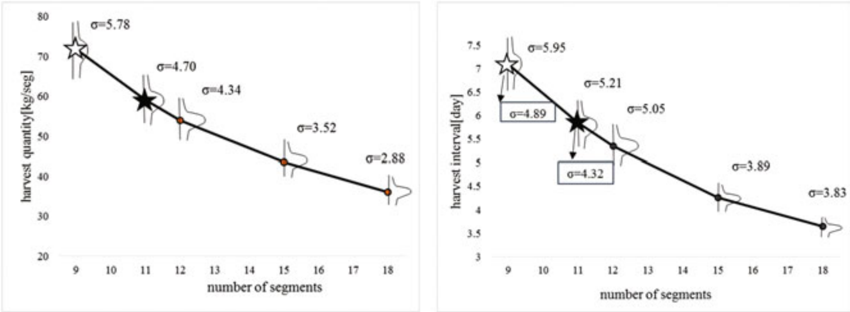


Fig. 7. Trial experiment results (five years)

6 Concluding Remarks

In this study, agricultural production management method is suggested for achieving Heijunka operation and applied to a production management division of collaborating agri-company. The turbulent factors in the agricultural production process were clarified and its performance is analyzed by simulation experiments by controlling levels of these factors. The obtained results revealed that elimination of turbulent elements such as seeding interval, growth duration and aggregated yield are critical for performance improvement. Furthermore seeding schedule on segmented farmland can respond flexibly to demand changes under restricted usability of farmland. Obtained result can be a fine example of lean management transfer to agriculture industry.

References

1. Chiyoma H, Murata K, Katayama H (2014) On performance analysis of sustainable closed logistics system in agricultural industry. In: Proceedings of the 17th annual conference of Japan society of logistics systems, pp 101–106
2. Hatake Company Co., Ltd. HP (2016) (in Japanese). http://www.t-k-f.jp/company.html
3. Ito K (2006) Mathematical programming problems in agricultural management. Oper Res Manage Sci 5:264–267 (in Japanese)
4. Katayama H (2016) Transfer activities of lean management to other industries-transplanting Heijunka concept for leanised operations. In: Proceedings of the 20th Cambridge international manufacturing symposium, pp 1–17
5. Katayama H, Aoki R (2016) On Heijanka manufacturing system for agricultural products. In: Proceedings of the 46th computers and industrial engineering, p 8
6. Ministry of Agriculture, Forestry and Fisheries HP (2015). http://www.maff.go.jp/
7. Vassian HJ (1955) Application of discrete variable servo theory to inventory control. J Oper Res Soc Am 3:272–282

Optimizing Reserve Combination with Uncertain Parameters (Case Study: Football)

Masoud Amel Monirian(✉), Hamideh Razavi, and Shabnam Mahmoudzadeh Vaziri

Industrial Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran
masoudmonirian@yahoo.com

Abstract. Preventive maintenance is critical when non-stop operation with maximum efficiency and reliability is required. In such conditions, proper identification of reserve and spare parts in terms of types and models, order and stock sizes as well as replacement timings become highly important. Arranging reserve sets, many constraints such as costs, space, obsolescence and deterioration must be met. Therefore, an integrated model for appropriate selection of reserves and timing for substitution is needed. This model can be similarly used in multiplayer sports such as football games. In this paper we propose an uncertain mathematical model to solve reserve combination strategy.

Keywords: Reserve combination · Optimization · Uncertainty theory · Uncertain programming

1 Introduction

Selecting an optimum reserve set and implementing the best replacement strategy is a challenge for most industries. Likewise, in multiplayer sports, choosing the right combination of the reserve team and taking a proper replacement strategy needs an advance planning and optimization. Inaccurate arrangements can cause significant financial loss as well as reputations.

When a team enters a tough league to play against different opponents with different methods, these issues become even more critical. Success or failure of each team depends on member's ability and skills. Therefore, it is essential to arrange the best combination both in main and reserve team in order to obtain the optimum results [2,9]. A few structured models have addressed this problem. Most of the studies emphasis on combination of the main team [1]. Since decisions are made prior to the games, we face probability and uncertainty in the problem, for which we use uncertain programming. It is a powerful tool in modeling industrial, human, natural and peddling and probability systems.

Certainly in intensive games, it is impossible to use 11 players. Players due to fatigue, injury, disciplinary cases and other reasons aren't available in all

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_119

games, and coach should use all main and reserve players to reach victory and profitability. Importance of reserve players when they replaced in proper time with suitable person, is not hidden to anyone, because with their adequate energy and by using skills and power can be influential. For instance, in final game of world cup 2014 in extra times, substitution players led the German team won the game.

1.1 Football Overview

Players are divided to 5 posts: goal keeper, defender, midfielder, sides, forward. According to game's technique, number of players in each post (expect goal keeper) can be different. This classification is done according to each player's abilities. Players due to their skills can be suitable for several posts.

1.2 Problem Definition

Broadly speaking, in selecting reserve team we assume that we have a main team and according to its constraints, situation and game circumstances, a reserve team should be selected. Replacement strategy conforming the game process and player's situation, declares that which player in the pitch should replace with which player in reserve team, in best time in order to preserve quality of the team. A football team plays in diverse leagues. Sometimes the schedule becomes so compressed, even they may have 3 or 4 games in a week. This event causes player's exhaustion. Consequently, generally main combination are not considered fix, so reserve combination selection problem and also replacement planning are very important issues. The first and main step for selecting reserve team and replacement strategy is recognition and evaluation of main team and desired options for reserve so that able us to compare different players in each post with stable criteria. First, players evaluated with belief degrees and by using mathematical model, choose reserve team to maintain team quality at maximum state. For rating players we need to consider several number of quantitative and qualitative characteristics in player's selection, like individual skills, statistics data in previous matches, physical readiness, psychological factors, injury and also opponent's condition [8]. In average, each football team, in normal situation, has more than 24 players that 10 of them plays in main combination and 6 of them are in reserve team. In this problem, we do not consider goal keeper selection. Because they run less during the game so their fatigue is less, but for compulsive situations, injury or expulsion it is necessary to have a goalkeeper in reserve team [3,8]. For scoring and evaluating players, 18 indexes and criteria such as heading, jumping, dribbling, tackling and so on are defined, that these criteria are approved by football simulation's companies. We value these criteria according to coach's opinion. In Table 1 these criteria are presented.

Uncertainty theory was founded by Liu [5] in 2007 and studied by many researchers. Uncertainty theory is a new method for modeling undetermined phenomena based on normality, duality, subadditivity and product axioms.

Table 1. List of criteria

Criteria	Number	Criteria	Number
Heading, jumping	1	Create goal scoring position	10
Shoot	2	Tackling	11
Short passing	3	Both feet	12
Crossing	4	Great stamina	13
Ball control	5	Height	14
Dribbling	6	Providing through (long) pass	15
Finishing (composure)	7	Technical ability	16
Speed	8	Create attacking opportunities	17
Creativity	9	Read the game	18

Uncertain programming presented by Liu [3], is a type of mathematical programming which involves uncertain variables. This paper aims to model uncertainty in football decisions based on uncertainty theory.

The rest of this paper is organized as follows. In Sect. 2, we review some basic concepts about uncertainty theory. Then we introduce uncertain programming in Sect. 3. After that, we introduce determining the reserve combination problem and uncertain determining the reserve combination problem in Sects. 4 and 5, respectively. In Sect. 6, we propose experimental results and solve the model. At last, some remarks are made in Sect. 7.

2 Preliminary

Uncertainty theory is a branch of axiomatic mathematics based on normality, duality, subadditivity and product axioms. In this section, we introduce some concepts in uncertainty theory, which are used throughout this paper.

Definition 1 (Liu [7])**.** Let τ be a nonempty set, and $\mathcal{L}$ be a $\sigma-$ algebra on Γ. A set function M is called an uncertain measure if it satisfies the following axioms,

Axiom 1: (Normality Axiom) $M\{\Gamma\} = 1$;

Axiom 2: (Duality Axiom) $M\{\Lambda\} + M\{\Lambda^c\} = 1$ for any event Λ;

Axiom 3: (Subadditivity Axiom) For every sequence of, we have

$$M\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \leq \sum_{i=1}^{\infty} M\{\Lambda_i\}. \tag{1}$$

In this case, the triple $(\Gamma, \mathcal{L}, M)$ is called an uncertainty space.

In addition, a product axiom on the product of uncertainty spaces was defined by Liu [7], which is the fourth axiom of uncertainty theory.

Axiom 4: (Product Axiom) Let $(\Gamma_k, \pounds_k, M_k)$ be uncertainty spaces for Then the product uncertain measure M is an uncertain measure satisfying

$$M\left\{\prod_{k=1}^{\infty} \Lambda_k\right\} = \Lambda_{k=1}^{\infty} M_k \left\{\Lambda_k\right\}, \tag{2}$$

where Λ_k are arbitrarily chosen events from $\pounds_k$ for $k = 1, 2, \cdots$, respectively.

Definition 2 (Liu [7]). An uncertain variable ξ is a measurable function from an uncertainty space $(\Gamma, \pounds, M)$ to the set of real numbers, i.e., for any Borel set ξ of real numbers, the set

$$\{\xi \in B\} = \{\gamma \in \Gamma | \xi(\gamma) \in B\} \tag{3}$$

is an event.

Definition 3 (Liu [4]). The uncertain variables $\xi_1, \xi_2, \cdots, \xi_n$ are said to be independent if

$$M\left\{\bigcap_{i=1}^{n} (\xi_i \in B_i)\right\} = \Lambda_{i=1}^{n} M\left\{\xi_i \in B_i\right\} \tag{4}$$

for any Borel sets $B_1, B_2, \cdots, B_n$ of real numbers.

Definition 4 (Liu [7]). The uncertainty distribution ϕ of an uncertain variable ξ is defined by

$$\phi(x) = M\left\{\xi \leq x\right\} \tag{5}$$

for any real number x.

If the uncertainty distribution $\phi(x)$ has an inverse function $\phi^{-1}(\alpha)$ for each $\alpha \in (0, 1)$, then $\phi^{-1}(\alpha)$ is called the inverse uncertainty distribution of the uncertain variable ξ. In this case, the uncertainty distribution ϕ is said to be regular. An uncertain variable ξ is called linear if it has an uncertainty distribution

$$\phi(x) = \begin{cases} 0, & \text{if}\, x < a \\ \frac{(x-a)}{(b-a)}, & \text{if}\, a \leq x \leq b \\ 1, & \text{if}\, x > b \end{cases} \tag{6}$$

denoted by $\pounds(a, b)$.

Theorem 1 (Liu [6]). *Let $\xi_1, \xi_2, \cdots ., \xi_n$ be independent uncertain variables with regular uncertainty distributions $\phi_1, \phi_2, \cdots, \phi_n$, respectively. If the function $f(x_1, x_2, \cdots, x_n)$ is strictly increasing with respect to $x_1, x_2, \cdots, x_m$ and strictly decreasing with respect to $x_{m+1}, x_{m+2}, \cdots, x_n$, then $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is an uncertain variable with an inverse uncertainty distribution*

$$\phi^{-1}(\alpha) = f(\phi_1^{-1}(\alpha), \cdots, \phi_m^{-1}(\alpha), \phi_{m+1}^{-1}(1-\alpha), \cdots, \phi_n^{-1}(1-\alpha)). \tag{7}$$

Definition 5 (Liu [7]). The expected value of an uncertain variable ξ is defined by

$$E[\xi] = \int_0^{+\infty} M\{\xi \geq x\}\,\mathrm{d}x - \int_{-\infty}^{0} M\{\xi \leq x\}\,\mathrm{d}x, \tag{8}$$

provided that at least one of the two integrals is finite.

If an uncertain variable ξ has a regular uncertainty distribution ϕ, then

$$E[\xi] = \int_0^1 \phi^{-1}(\alpha)\mathrm{d}\alpha \tag{9}$$

provided that the expected value $E[\xi]$ exists. Lease acknowledge collaborators or anyone who has helped with the paper at the end of the text.

3 Uncertain Programming

Uncertain programming, which was first proposed by Liu [5], is a type of mathematical programming involving uncertain variables. Assume that x is a decision vector, and ξ is an uncertain vector. Since the uncertain programming model contains the uncertain objective function $f(x,\xi)$ and uncertain constraints $g_j(x,\xi) \leq 0$, $j = 1, 2, \cdots, p$, Liu [5] proposed the following uncertain programming model,

$$\begin{cases} \min_x & E[f(x,\xi)] \\ \text{s. t.} & M\{g_j(x,\xi) \leq 0\} \geq \alpha_j, \quad j = 1, 2, \cdots, p. \end{cases} \tag{10}$$

Definition 6 (Liu [6]). A vector is called a feasible solution to the uncertain programming model if

$$M\{g_j(x,\xi) \leq 0\} \geq \alpha_j \tag{11}$$

for $j = 1, 2, \cdots, p$.

Definition 7 (Liu [6]). A feasible solution x^* is called an optimal solution to the uncertain programming model if

$$E[f(x^*,\xi)] \leq E[f(x,\xi)] \tag{12}$$

for any feasible solution x.

Theorem 2 (Liu [6]). *Assume $f(x, \xi_1, \xi_2, \cdots, \xi_n)$ is strictly increasing with respect to $\xi_1, \xi_2, \cdots, \xi_m$ and strictly decreasing with respect to $\xi_{m+1}, \xi_{m+2}, \cdots, \xi_n$, and $g_j(x, \xi_1, \xi_2, \cdots, \xi_n)$ are strictly increasing with respect to $\xi_1, \xi_2, \cdots, \xi_k$ and strictly decreasing with respect to $\xi_{k+1}, \xi_{k+2}, \cdots, \xi_n$ for $j = 1, 2, \cdots, p$.*

If $\xi_1, \xi_2, \cdots, \xi_n$ *are independent uncertain variables with uncertainty distributions* $\phi_1, \phi_2, \cdots, \phi_n$*, respectively, then the uncertain programming*

$$\begin{cases} \min_x & E\left[f(x, \xi_1, \xi_2, \cdots, \xi_n)\right] \\ s.\ t. & M\left\{g_j(x, \xi_1, \xi_2, \cdots, \xi_n) \leq 0\right\} \geq \alpha_j, \quad j = 1, 2, \cdots, p. \end{cases} \tag{13}$$

is equivalent to the crisp mathematical programming

$$\begin{cases} \min_x & \int_0^1 f(x, \phi_1^{-1}(\alpha), \cdots, \phi_m^{-1}(\alpha), \phi_{m+1}^{-1}(1-\alpha), \cdots, \phi_n^{-1}(1-\alpha))\mathrm{d}\alpha \\ s.\ t. & g_j(x, \phi_1^{-1}(\alpha_j), \cdots, \phi_k^{-1}(\alpha_j), \phi_{k+1}^{-1}(1-\alpha_j), \cdots, \phi_n^{-1}(1-\alpha_j)) \leq 0, \\ & j = 1, 2, \cdots, p. \end{cases} \tag{14}$$

4 Mathematical Model for Determining the Reserve Combination

In this model, due to the evaluations from players on the pitch and reserve players and considering constraints, number of players are chosen to be in reserve team, aims to maximize the power of reserve team and minimize the reduction of team's weakness. The model is linear with binary variables.
Indices:

n: number of players (in football it is usually 22 players);
k: number of features $k = 1, 2, \cdots, 18$ (Table 2);
i: the number of fixed players $i = 1, 2, \cdots, 10$;
j: the number of reserve players $j = 1, \cdots, n - i$;
l: number of posts $l = 1, \cdots, 4$ (for values of l we have:
$l = 1$ defender, $l = 2$ midfielder, $l = 3$ sides, $l = 4$ forward).

Parameters:

c_{jk}: the value of criteria k of player;
$MinR$: minimum number of reserve players;
w_{kl}: the value of criteria k in post l;
$MaxR$: maximum number of reserve players;
L_i: the post of player i;
P_i: probability of need to replace player i;
R_j: player accessibility characteristic (if player j is available $R_j = 1$, otherwise $R_j = 0$).

Decision variable:

$$x_{ij} = \begin{cases} 1, \text{ if there is need to replace player } i \text{ with player } j \\ 0, \text{ otherwise.} \end{cases}$$

The model is defined as follows,

$$\max \sum_{k} \sum_{i} \sum_{j} w_{kL_i}\, c_{jk}\, p_i\, x_{ij} \tag{15}$$

$$\text{s.t} \sum_{i} \sum_{j} x_{ij} \leq \max R \tag{16}$$

$$\sum_{i} x_{ij} \leq R_j \quad \forall j \tag{17}$$

$$\sum_{j} x_{ij} \geq 1 \quad \forall i,\ \mathrm{L_i} = 1, \quad \forall i, \quad \mathrm{L_i} = 2, \quad \forall i, \quad \mathrm{L_i} = 3, \quad \forall i, \quad \mathrm{L_i} = 4 \tag{18}$$

$$x_{ij} \in \{0, 1\}. \tag{19}$$

The proposed model is linear. The objective function (15) targets to maximize the reserve team's quality, which obtains from multiplying each player's value in desired post in probability of the need to replacement in that post. Constraint (16) determines maximum number of allowable replacements. Constraint (17) indicates the availability of players because it is possible that the player is not in the team owing to Injury or deprivation of the game. Constraint (18) states that each post should not be vacant. Constraint (19) indicates decision variable type.

5 Uncertain Mathematical Model for Determining the Reserve Combination

The player's criteria are not constant always and they can change due to game situation, fatigue and so on. To use the criteria in mathematical model, we use belief degree and uncertain programming.

The parameters c_{jk}, w_{kL_i} and p_i are uncertain parameters with uncertain distributions $\varphi_{jk}, \phi_{kL_i}$ and ω_i, respectively. We define $f(x, C, W, P) = \sum_k \sum_i \sum_j C_{jk} W_{kL_i} P_i x_{ij}$, whose inverse uncertainty distribution is $\gamma^{-1}(x, \alpha) = \sum_k \sum_i \sum_j \varphi_{jk}^{-1} \phi_{kL_i}^{-1} \omega_i^{-1} x_{ij}$.

To use uncertain programming the model changes as follows, Since $\gamma^{-1}(x, \alpha)$ is the inverse uncertainty distribution of $f(x, C, W, P)$, the mathematical model is simplified as follows,

$$\max E\left[f(x, C, W, P)\right] \tag{20}$$

$$\text{s.t.} \sum_{i} \sum_{j} x_{ij} \leq MaxR \tag{21}$$

$$\sum_{i} x_{ij} \leq R_j \quad \forall j \tag{22}$$

$$\sum_{j} x_{ij} \geq 1 \quad \forall i,\ \mathrm{L_i} = 1, \forall i, \quad \mathrm{L_i} = 2, \quad \forall i, \quad \mathrm{L_i} = 3,\ \forall i, \quad \mathrm{L_i} = 4 \tag{23}$$

$$x_{ij} \in \{0, 1\}. \tag{24}$$

6 Experimental Results

The proposed model solved with IBM ILOG CPLEX-Optimizer v12. 3. We choose a team with 24 players and solve the problem with these players. Due to coach's opinion the belief degrees are obtained. The results are presented in appendix. In football games the maximum allowed replacements as 3 times. The results of solving models using these data are examined in this section. The main players' number, which is determined according to coach's opinion, is like follows. This will be an input for model. The reserve team will be chosen due to the input of model.

22	20	18	15	14	9	5	2	7	1

The reserve team, due to the model, is chosen as follows,

12	23	19	11	8	3

According to model suggested combination, average quality of the reserve will be 83%. When only the coach's opinion was important, and without using model, the average quality of the reserve was 55% that brings a significant increase for team. The reserve team, according to coach's opinion, shows as follows,

10	24	19	11	4	3

Table 2 shows the comparison of objective functions in two cases which indicates the model's efficiency.

Table 2. Comparison of the first model in simulation environment

Number of simulation runs	$\overline{f^*}$ without the model	$\overline{f^*}$ by the model	Improvement of simulation runs
80	63%	91%	28%

7 Conclusion

The purpose of this article is to propose a model for choosing reserve team, so the quality reduction will be minimum. We proposed a linear mathematical model. To obtain uncertain parameters we use belief degree and uncertain programming. For future works, authors can investigate main combination selection due to the future games and simultaneous select reserve team.

Acknowledgement. This work was granted from Professor Baoding Liu, Tsinghua University. The authors appreciate his financial and scientific supports. We also appreciate Professor Mitsuo Gen for his guidance.

Appendix

See Tables 3, 4, 5, 6, and 7.

Table 3. The belief degrees of coach on all criteria

Number of criteria	Defender	Midfielder	Sides	Forwards	Number of criteria	Defender	Midfielder	Sides	Forwards
C1	(0.9,1)	(0.4,0.5)	(0.6,0.75)	(0.85,1)	C10	(0.45,0.55)	(0.6,0.75)	(0.6,0.75)	(0.85,1)
C2	(0.45,0.55)	(0.6,0.75)	(0.6,0.75)	(0.85,1)	C11	(0.85,1)	(0.6,0.75)	(0.45,0.55)	(0.85,1)
C3	(0.85,1)	(0.85,1)	(0.85,1)	(0.6,0.75)	C12	(0.45,0.55)	(0.45,0.85)	(0.45,0.55)	(0.6,0.75)
C4	(0.45,0.55)	(0.6,0.75)	(0.85,1)	(0.6,0.75)	C13	(0.4,0.5)	(0.85,1)	(0.4,0.5)	(0.6,0.75)
C5	(0.6,0.75)	(0.85,1)	(0.85,1)	(0.85,1)	C14	(0.6,0.75)	(0.45,0.55)	(0.85,1)	(0.85,1)
C6	(0.45,0.55)	(0.9,1)	(0.6,0.75)	(0.85,1)	C15	(0.3,0.45)	(0.85,1)	(0.6,0.75)	(0.6,0.75)
C7	(0.45,0.55)	(0.4,0.5)	(0.6,0.75)	(0.85,1)	C16	(0.3,0.45)	(0.6,0.75)	(0.85,1)	(0.85,1)
C8	(0.45,0.55)	(0.6,0.75)	(0.85,1)	(0.85,1)	C17	(0.6,0.75)	(0.85,1)	(0.6,0.75)	(0.85,1)
C9	(0.45,0.55)	(0.85,1)	(0.9,1)	(0.6,0.75)	C18	(0.6,0.75)	(0.85,1)	(0.85,1)	(0.6,0.75)

Table 4. The belief degrees of coach on all player

Number of players	1	2	3	4	5	6	7	8
P	(0.6,0.75)	(0.3,0.5)	(0.1,0.3)	(0.1,0.2)	(0.3,0.5)	(0.3,0.5)	(0.1,0.3)	(0.3,0.4)
Number of players	9	10	11	12	13	14	15	16
P	(0.4,0.5)	(0.5,0.6)	(0.3,0.4)	(0.55,0.8)	(0.2,0.3)	(0.3,0.5)	(0.6,0.8)	(0.3,0.45)
Number of players	17	18	19	20	21	22	23	24
P	(0.3,0.4)	(0.5,0.7)	(0.1,0.3)	(0.1,0.3)	(0.3,0.5)	(0.3,0.45)	(0.3,0.5)	(0.3,0.4)

Table 5. The belief degrees of coach on all criteria of different players

Number of criteria	Number of players							
	Defender							
	1	2	3	4	5	6	7	8
C1	(0.7,0.8)	(0.7,0.8)	(0.8,0.95)	(0.6,0.7)	(0.8,0.95)	(0.45,0.6)	(0.95,1)	(0.7,0.8)
C2	(0.45,0.6)	(0.95,1)	(0.8,0.95)	(0.25,0.4)	(0.35,0.5)	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)
C3	(0.7,0.8)	(0.7,0.8)	(0.8,0.95)	(0.7,0.8)	(0.8,0.95)	(0.35,0.5)	(0.5,0.6)	(0.7,0.8)
C4	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.25,0.4)	(0.7,0.8)	(0.95,1)	(0.45,0.6)	(0.7,0.8)
C5	(0.7,0.8)	(0.45,0.6)	(0.45,0.6)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)
C6	(0.45,0.6)	(0.7,0.8)	(0.7,0.85)	(0.7,0.8)	(0.6,0.7)	(0.6,0.7)	(0.25,0.4)	(0.7,0.8)
C7	(0.45,0.6)	(0.6,0.7)	(0.6,0.7)	(0.25,0.4)	(0.6,0.7)	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)
C8	(0.7,0.8)	(0.45,0.6)	(0.35,0.5)	(0.8,0.95)	(0.5,0.6)	(0.6,0.7)	(0.7,0.8)	(0.95,1)
C9	(0.6,0.7)	(0.45,0.6)	(0.5,0.6)	(0.6,0.7)	(0.25,0.4)	(0.7,0.8)	(0.7,0.8)	(0.95,1)
C10	(0.45,0.6)	(0.7,0.8)	(0.25,0.4)	(0.45,0.6)	(0.25,0.4)	(0.45,0.6)	(0.45,0.6)	(0.7,0.8)
C11	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)
C12	(0.25,0.4)	(0.7,0.8)	(0.25,0.4)	(0.45,0.6)	(0.45,0.6)	(0.7,0.8)	(0.25,0.4)	(0.6,0.7)
C13	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.7,0.8)	(0.7,0.8)
C14	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.95,1)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)
C15	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.45,0.6)	(0.45,0.6)	(0.45,0.6)
C16	(0.7,0.8)	(0.25,0.4)	(0.45,0.6)	(0.6,0.7)	(0.45,0.6)	(0.6,0.7)	(0.5,0.6)	(0.6,0.7)
C17	(0.45,0.6)	(0.7,0.8)	(0.45,0.6)	(0.25,0.4)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.95,1)
C18	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.7,0.8)	(0.6,0.7)

Table 6. The belief degrees of coach on all criteria of different players

Number of criteria	Number of players						
	Midfielder						
	9	10	11	12	13	14	15
C1	(0.7,0.8)	(0.35,0.5)	(0.7,0.8)	(0.45,0.6)	(0.25,0.4)	(0.6,0.7)	(0.95,1)
C2	(0.8,0.95)	(0.7,0.8)	(0.95,1)	(0.35,0.5)	(0.7,0.8)	(0.95,1)	(0.7,0.8)
C3	(0.8,0.95)	(0.7,0.85)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)
C4	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.7,0.8)
C5	(0.95,1)	(0.6,0.7)	(0.8,0.95)	(0.7,0.8)	(0.45,0.6)	(0.5,0.6)	(0.45,0.6)
C6	(0.7,0.8)	(0.45,0.6)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.8,0.95)
C7	(0.45,0.6)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.95,1)
C8	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.45,0.6)	(0.95,1)	(0.95,1)	(0.7,0.8)
C9	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.95,1)	(0.45,0.6)	(0.7,0.8)
C10	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.7,0.8)
C11	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.45,0.6)	(0.45,0.6)	(0.7,0.8)
C12	(0.6,0.7)	(0.5,0.6)	(0.5,0.6)	(0.6,0.7)	(0.6,0.7)	(0.45,0.6)	(0.7,0.8)
C13	(0.7,0.85)	(0.6,0.7)	(0.6,0.7)	(0.6,0.7)	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)
C14	(0.45,0.6)	(0.45,0.6)	(0.7,0.8)	(0.45,0.6)	(0.8,0.95)	(0.6,0.7)	(0.6,0.7)
C15	(0.6,0.7)	(0.8,0.95)	(0.6,0.7)	(0.6,0.7)	(0.8,0.95)	(0.7,0.85)	(0.6,0.7)
C16	(0.7,0.8)	(0.6,0.7)	(0.6,0.7)	(0.7,0.8)	(0.7,0.85)	(0.95,1)	(0.7,0.8)
C17	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.7,0.8)
C18	(0.8,0.95)	(0.95,1)	(0.45,0.6)	(0.7,0.85)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)

Table 7. The belief degrees of coach on all criteria of different players

Number of criteria	Number of players								
	Sides				Forward				
	16	17	18	19	20	21	22	23	24
C1	(0.7,0.8)	(0.7,0.8)	(0.8,0.95)	(0.7,0.8)	(0.95,1)	(0.8,0.95)	(0.95,1)	(0.95,1)	(0.7,0.8)
C2	(0.95,1)	(0.95,1)	(0.8,0.95)	(0.7,0.8)	(0.95,1)	(0.8,0.95)	(0.6,0.7)	(0.7,0.8)	(0.7,0.8)
C3	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.95,1)
C4	(0.45,0.6)	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.7,0.8)
C5	(0.95,1)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.8,0.95)	(0.6,0.7)	(0.7,0.8)	(0.7,0.8)
C6	(0.7,0.85)	(0.7,0.8)	(0.95,1)	(0.6,0.7)	(0.8,0.95)	(0.6,0.7)	(0.7,0.8)	(0.7,0.8)	(0.95,1)
C7	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.95,1)
C8	(0.45,0.6)	(0.95,1)	(0.7,0.8)	(0.8,0.95)	(0.95,1)	(0.7,0.8)	(0.8,0.95)	(0.8,0.95)	(0.7,0.8)
C9	(0.95,1)	(0.95,1)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.95,1)
C10	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.45,0.6)	(0.95,1)	(0.45,0.6)
C11	(0.95,1)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)
C12	(0.25,0.4)	(0.7,0.8)	(0.7,0.8)	(0.8,0.95)	(0.7,0.85)	(0.7,0.85)	(0.5,0.6)	(0.7,0.85)	(0.6,0.7)
C13	(0.7,0.8)	(0.45,0.6)	(0.6,0.7)	(0.8,0.95)	(0.8,0.95)	(0.7,0.85)	(0.7,0.8)	(0.95,1)	(0.7,0.8)
C14	(0.8,0.95)	(0.95,1)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.7,0.8)
C15	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)
C16	(0.7,0.85)	(0.7,0.8)	(0.8,0.95)	(0.7,0.8)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)	(0.95,1)
C17	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.7,0.8)	(0.45,0.6)	(0.95,1)	(0.7,0.8)	(0.7,0.8)	(0.95,1)
C18	(0.95,1)	(0.45,0.6)	(0.7,0.8)	(0.45,0.6)	(0.7,0.8)	(0.45,0.6)	(0.95,1)	(0.85,0.95)	(0.7,0.8)

References

1. Abdulwhab A, Billinton R et al (2004) Maintenance scheduling optimization using a genetic algorithm (GA) with a probabilistic fitness function. Electr Power Compon Syst 32(12):1239–1254
2. Ahmed F, Deb K, Jindal A (2013) Multi-objective optimization and decision making approaches to cricket team selection. Appl Soft Comput 13(1):402–414
3. Liu B (2002) Theory and Practice of Uncertain Programming. Physica-Verlag, Heidelberg
4. Liu B (2009) Some research problems in uncertainty theory. J Uncertain Syst 3(1):3–10
5. Liu B (2009) Uncertain entailment and modus ponens in the framework of uncertain logic. J Uncertain Syst 3(4):243–251
6. Liu B (2010) Uncertainty Theory: A Branch of Mathematics for Modeling Human Uncertainty. DBLP
7. Liu B (2015) Uncertainty theory. Stud Comput Intell 154(3):1–79
8. Tavana M, Azizi F et al (2013) A fuzzy inference system with application to player selection and team formation in multi-player sports. Sport Manag Rev 16(1):97–110
9. Trninić S, Papić V et al (2008) Player selection procedures in team sports games. Acta Kinesiologica 2(1):24–28

A Bayesian-Based Co-Cooperative Particle Swarm Optimization for Flexible Manufacturing System Under Stochastic Environment

Lu Sun[1], Lin Lin[1,2,3(✉)], and Haojie Li[1]

[1] Dalian University of Technology, Dalian 116024, People's Republic of China
lin@dlut.edu.cn
[2] Fuzzy Logic Systems Institute, Tokyo, Japan
[3] Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian, People's Republic of China

Abstract. In recent years, flexible scheduling attracted considerable attention, motivated by both important practical issues and interesting research problems, especially under stochastic environment. In this paper we discuss a stochastic scheduling problem whose scale arranges from small to large. The objective in the schedule is to minimize the processing time over all of the jobs. However, stochastic environment will bring more difficulty especially for large scale because the sudden change of processing time for each operation may break the current optimal solution and lose efficiency. So, we propose a Bayesian network based particle swarm optimization (BNPSO) for solving this stochastic scheduling problem. Firstly, we use new framework named co-cooperative (CC) evolutionary framework which decompose all decision variables into several small group containing part of decision variables to overcome large scale problems. And then, BNPSO adjust the group scene according to their interactive relationships based on Bayesian network structure. Meanwhile, importing self-adaptive mechanism for parameters in order to satisfy the stochastic environment. Some practical test instances will demonstrate the effectiveness and efficiency of the proposed algorithm.

Keywords: Co-cooperative framework · Bayesian network structure · Stochastic environment · Flexible manufacturing system

1 Introduction

Over the past sixty years, a great number of researches have been conducted on the job shop scheduling problem (JSP), which is a branch of the scheduling problem and highly popular in the manufacturing industry. JSP is a classical combinatorial optimization problem, and it is recognized as NP-hard under the precedence and resource constraints [2,5]. The flexible job shop scheduling problem (FJSP) is a generalization of the classical JSP for flexible manufacturing systems (FMSs).

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_120

In real-world problem, most of the scheduling problems are the stochastic scheduling problems. As one of the newest issues, more and more attentions focused on the problem with random processing time. As a result, in the last several decades, a significant amount of results has been achieved on the stochastic scheduling problems [3]. Moghaddam et al. [13] proposed a hybrid method using a neural network approach and a simulated annealing algorithm in two stages, in order to produce the optimal/near-optimal solution. Liu, Wang and Jin [9] presented an approach named PSOSAHT that is hybrid with simulated annealing (SA) and hypothesis test (HT), for stochastic flow shop scheduling with uncertain processing time. Zhou, Nee and Lee [17] developed an ant colony optimization algorithm (ACO) with different levels of machine utilizations, processing time distributions, and performance measures. Lei [6] developed an efficient decomposition-integration genetic algorithm (DIGA) to minimize the maximum fuzzy completion time. For solving realistic scheduling problems, intelligent manufacturing planning and scheduling based on meta-heuristics, such as GAs and ACOs have become common techniques for finding satisfactory solutions within reasonable computational times. However, as the scale of problem increases, the search space will increase exponentially, consequently and the probability of obtaining optimal solution will be reduced. To overcome it, coevolution framework was proposed firstly in 1994, which represents and solves more complex problems by explicitly modeling the coevolution of cooperating species [11]. Up to now, there exist various typical grouping mechanisms. Bergh et al. [14] proposed a cooperative coevolutionary approach which used the fixed group size to solve function problems. Li, et al. [10] thought more frequent random grouping mechanism could have better performance and proposed a cooperative coevolutionary framework and employs a technique called random grouping in order to group interacting variables into one subcomponent for large scale continuous global function optimization. Yao et al. [8] balanced the random grouping with fixed grouping, and proposed a set-based grouping mechanism which group set was given in advanced. When the global optimal solution does not be optimized, to select a new group size randomly immediately. However, all of the grouping strategies described so far use a predefined and fixed group size. For example, random grouping decomposes an n-dimensional problem into k s-dimensional problems. A major drawback of these techniques is that the user needs to specify a value for either k or s or a set of potential s values and the given values will effect the performance directly. Another drawback of this multilevel scheme is that once an s value is chosen, the decision variables are divided into a set of equally sized subcomponents. It is unlikely that in most real-world problems the sizes of interacting groups will be not equal. Hence, it is desirable that a decomposition strategy can automatically determine the number of subcomponents and their sizes.

In recent years, probabilistic graphical models (PGMs), which are used to represent the relationship among the discrete variables, have attracted a growing attention. PGMs can be mainly categorized into two categories: Bayesian networks (directed) and Markov networks (undirected) [15]. In Bayesian network, nodes represent variables and arcs shows the relationship between two linked nodes. So, if we can use Bayesian network to shown the relationships among the variables, we will no longer need to determine the number of subcomponents and their sizes. In the application of Bayesian network, Bayesian optimization algorithm (BOA) is widely used in scheduling, such as, resource assignment, task scheduling and so on. In 2011, Yang et al. [15] present a novel scheduling algorithm based on BOA for heterogeneous computing environments. Li et al. [7] proposed a Bayesian optimization algorithm to solve task assignment problems in heterogeneous computing systems. Hao et al. [4] proposed a novel cooperative Bayesian optimization algorithm (CoBOA) to overcome the challenges mentioned in the field of multiple resources scheduling problem (MRSP). Inspired by BOA, in this paper, we proposed a learning-based grouping mechanism in which we apply BOA to get the optimal Bayesian network structure for finding the potential relationships among variables and then to divide the variables according to relationships (network structure).

The remainder of this paper is organized as follows: Sect. 2 gives the formulating process of S-fJSP and mathematical programming model; Sect. 3 describes the BNPSO combined BOA Particle swarm algorithm (PSO) with multiple subpopulations in detail; Sect. 4 gives kinds of detailed numerical experiments and computational results; finally, Sect. 5 comes to the conclusion of the paper.

2 Mathematical Formulation

In order to solve a stochastic flexible job-shop scheduling problem (S-fJSP), it assumes that the probability distribution of the processing time is known in advance. In this paper, we use a pure integer programming model to transmute the processing times in terms of stochastic variable. The S-JSP can be formulated as an extended version of fJSP. Difference to the conventional fJSP, each operation o_{ij} is carried out under uncertain random disturbance with pre-given expected valued $E\left[p_{ij}\right]$ and variance v_{ij}, where p_{ij} is the processing time of O_{ij} on the machines. The distribution of the variance can be predicted from the experimental data such as normal distribution, uniform distribution and exponential distribution etc. It may also consist of several assumptions as follows:

- A1. Each machine processes only one job at a time.
- A2. Each job is processed on one machine at a time.
- A3. For each operation, the probability distribution of processing time and machine assignment are known in advance, but the processing time is stochastic.
- A4. There are no precedence constraints among operations of different jobs.
- A5. Operations cannot be interrupted.
- A6. Neither release times nor due dates are specified.

The S-fJSP has already been confirmed as one of the NP-hard combinatorial problems. There are N jobs and M machines to be scheduled; furthermore, each job is composed of a set of operations and the operation order on machines is pre-specified. Each operation is characterized by the required machine and the processing time [5].

Index:

i,h: indices of tasks, $i, h = 1, 2, \cdots, n$;
j: indices of resources, index, $j = 1, 2, \cdots, m$;
k,g: indices of operations, $k, g = 1, 2, \cdots, n_i$.

Parameter:

n: number of total tasks;
m: number of total resources;
n_i: number of operations for task i;
n_{ik}: k^{th} operation of task i;
t_{ikj}: processing time of O_{ik} processed on j^{th} resource;
A_{ik}: the available resource set of O_{ik};
c_{ik}: the completed time of O_{ik}.

Decision variables:

$$x_{ikj} = \begin{cases} 1, \text{ if } o_{ik} \text{ is performed on machine } j \\ 0, \text{ otherwise.} \end{cases}$$

The objective function is minimizing the makespan, as follows:

$$\min E\left[C_M\right] = E\left[\max_i \left[\max_k\right] \left[\max_i \xi c_{ikj}\right]\right]. \tag{1}$$

For the process of scheduling, the operations are not allowed to terminate until it is completed and for each resource. The resource constraints are described as follows in Eq. (2):

$$s.t. \quad \xi c_{ik} - \xi c_{i(k-1)} \geq \xi t_{ikj} \times x_{ikj} \qquad k = 2, \cdots, n_i, \forall i, j. \tag{2}$$

For each task, it consists of several operations and each operation is allocated different machine, although each operation can be processed on different resources, the operation sequence within each task must be observed. Note that the two constraints (3) and (4) only one need to be satisfied at the same time.

$$(\xi c_{hg} - \xi c_{i(k-1)} - \xi t_{ikj}) \times x_{hgj} \geq 0 \quad \forall j, g, h, \tag{3}$$

$$(\xi c_{ik} - \xi c_{hg} - \xi t_{ikj}) x_{ikj} \geq 0 \quad \forall i, j, g, h. \tag{4}$$

Besides the constraints mentioned above, the constraint (5) guarantees machine allocation that for each operation can only process on one machine from machine set at one time.

$$\sum_{x_{ijk} \in A_{ik}} x_{ikj} = 1 \quad \forall i, k, j. \tag{5}$$

For each resource, it contains only two cases: chosen or not chosen. And The constraint (6) and the constraint (7) gives the restriction constraint of decision variables.

$$x_{ikj} \in \{0,1\} \quad \forall i,k,j, \tag{6}$$

$$\xi c_{ik} > 0 \quad \forall i,k. \tag{7}$$

3 Algorithm Design

BNPSO starts by random grouping all decision variables encoded by real numbers into several subcomponents, each subcomponent contains s variables shown in Fig. 1. Then, to optimize the global optimal solution based on the co-evolutionary framework. In the process of evolution, to record down the data set for BN learning according to the value variation of each decision variable. Each 50 iterations, update the structure of BN follows BOA and adjust the grouping scene. Meanwhile, self-adjust the parameters used in evolution algorithms. Several key parts of BNPSO listed in the following subsections in detail.

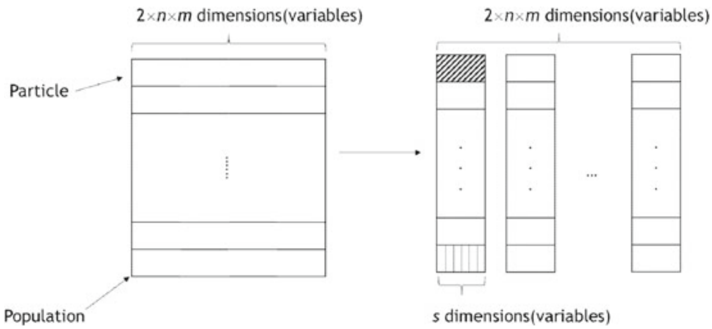


Fig. 1. The network based for decoding the operations

3.1 Encoding and Decoding

PSO is a typical met-heuristics algorithm in EAs, so the encoding and decoding methods will effect the performance partly. We choose two-section real priority encoding method in order to explore larger solution space. An example of representation and decoding network for one particle with 5 operations is shown Fig. 2.

The searching space is created in a $2 \times n \times m$ search space for n jobs on m machines. Each particle consists of $2 \times n \times m$ variables and is represented by real number. As mentioned above, S-fJSP can be considered as two subproblems: operation sequences and machine allocation. So the first section of particle presents the operations; the second part presents the machines allocation

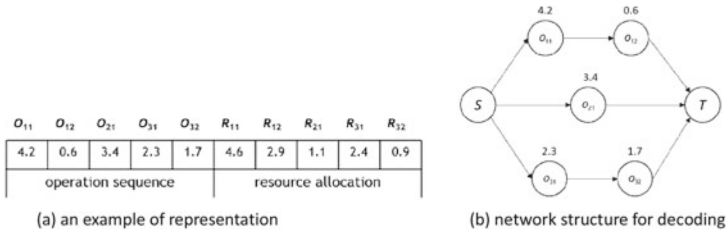


Fig. 2. An example of representation and decoding network for one particle

corresponding to each operation, we use the network based decoding method (shown in Fig. 3(b)) to get S-fJSP solution from the particle which can confirm the solutions are all feasible solutions.

PSO uses the following formulas to update velocity and position in each generation t, respectively. The initial values of velocity and position are set randomly with the lower bound and upon bound which are given in advance. The ω presents inertia weight, rand_1 and rand_2 are random value within [0,1]. The bigger ω means new velocity will mainly be effected by search history of particle itself; the smaller presents new velocity will mainly be effected by search history of the whole population. The parameters used in PSO is adapted according to the methods studied in the paper [10].

$$v_t(t+1) = \omega v_i(t) + c_1 \text{rand}_1[p_{\text{best}}(t) - x_{i\partial}(t)] + c_2 \text{rand}_2[l_{\text{best}}(t) - x_i(t)], \quad (8)$$

$$x_t(t+1) = x_t(t) + v_t(t+1). \quad (9)$$

In order to explore larger solution space, the velocity and position update model with both Cauchy and Gaussian distributions was proposed in [15]. This model updates the position directly through personal best: p_{best}, local best: l_{best}. p_{best} means the best particle in the searching history of each particle, l_{best} means the best one among the ith particle, $(i-1)^{th}$ particle and $(i+1)^{th}$ particle each iteration. rand and p are used for deciding which formula to be selected, rand is a random value within [0,1] and p is a given value at initial stage. $C(1)$ means the Cauchy distribution value with mean value 1; $N(0,1)$ means the normal distribution value with mean value 0 and the standard variance 1. The position updating formula is shown as follows, each time, generate rand, and compare rand with p to decide which equation to be chosen. If rand $\leq p$, the position will update follow the first formula, otherwise, follow the second formula.

$$x(t+1) = \begin{cases} p_{\text{best}}(t) + C(10\,|p_{\text{best}}(t) - l_{\text{best}}(t)|\,, & \text{if rand} \leq p \\ l_{\text{best}}(t) + N(0,1)\,|p_{\text{best}}(t) - l_{\text{best}}(t)|\,, & \text{otherwise.} \end{cases} \quad (10)$$

3.2 Bayesian-Based Grouping Mechanism

(1) Bayesian network structure learning
As mentioned above, BN structure learning is also a NP hard problem, if we use the traditional method to learn its structure, it will make our algorithm be too

heavy, so we use another evolutionary algorithm. Inspired by $K2$ which is the most classical and typical algorithm for learning BN structure, learning structure can be divided into two parts: the node sequence and the edges. Hence, we still use two-section encoding, in which the first section represents node sequence, and the second section represents the edge. We give an example with 4 variables in one group in Fig. 3(a). We can get variable sequence (descending order): $x_4- > x_2- > x_1- > x_3$, then get the edges by transferring the real values in the second part particle into binary value for the first 4 (equals to the number of variables s in current group) values, then to randomly select from the got binary values for other 2 (equals to $s \times (s-1)/2 - s$) binary values, shown in Fig. 3(b).

1	2	3	4	1	2	3	4
1.5	2.3	0.6	3.1	2.8	1.7	1.9	1.3
variable sequence				edge consideration			

(a) real number based encoding particle

x_4	x_2	x_1	x_3	1	0	0	1	0	0
variable sequence				edge consideration					

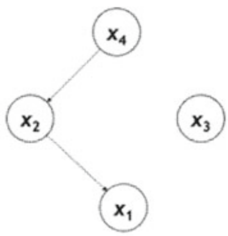
(b) particle after decoding

Fig. 3. The particle used for learning BN structure

Then, we decode particle into a BN by a strictly upper triangular matrix which means to put the binary sequence in the shadow position successively from the bottom to the end in Fig. 4. In which, 1 represents the variable in column is the parent of the variable in row. Each generation, we can get one network structure, after m generation, we can m candidate networks. A simple BN is shown in Fig. 4, it presents that the network constructed by particle in Fig. 3(a). Then, we adjust the grouping situation depends on BN structure, x_3 is a dependent variable, hence, we put it to another random group, after adjusting, this group only contain 3 variables temporary until other variables are put in.

	x_4	x_2	x_1	x_3
x_4		1	0	0
x_2			1	0
x_1				0
x_3				

(a) a strictly upper triangular matrix



(b) Bayesian network structure

Fig. 4. A simple BN structure example get from a particle

(2) Sampling train data
We record down the variance of each decision variable of particle between $(i+1)^{th}$ generation and the i^{th} generation. Because, the variance of each variable of global best particle shows the relationship when global best particle evaluates.

1	2	3	4
1.2	3.2	3.9	2.7
1.2	3.2	3.9	2.7
0.8	2.9	5.1	2.1
....			
1.0	2.7	3.0	2.3
1.2	2.5	2.9	2.7

(a) recorded global best particle

1	2	3	4
0	0	0	0
0	0	1	0
...			
1	0	0	1
1	0	0	1

(b) training data

Fig. 5. Sampling the training data

If the value of variable becomes bigger, then to record down 1, otherwise, if the value becomes smaller or keeps invariant, to record down 0. Then we can get the training data set for learning BN structure, shown in Fig. 5. After evaluating the network structure, we can find a optimal network with BDe score. Then, we adjust the group by putting the discrete variables into the other groups randomly. The whole process of BNPSO is shown as follows.

4 Experiment and Results

To prove the efficiency of our proposed algorithm logically, we do two kinds of experiments for S-fJSP on 4 scales of data: 5×5 (50 dims), 10×10 (200 dims), 15×15(450 dims) and 20×20 (800 dims). The dimensions are decided by encoding method mentioned in Sect. 3. For example, 5×5 means this benchmark has 5 jobs, 5 machines and 5 (the same to the number of machines) operations. These data instances can be accessed on OR-Library by Dirk et al. For the uncertainty processing time, the processing time of operation o_{ij} may equally take any real value from the uniform distribution U(lower, upon) where lower is 80% of static processing time and upon is 120% of static processing time. In order to ensure the reliability of the experiments and avoid the shock of results, all experiments are repeated 30 times to get the mean value and calculate variance. The machine environment of running experiments is Intel(R) Core(TM) i7-4770 CPU @ 2.2 GHz, 16 GB.

4.1 Experiment Settings

All compared algorithms contains: classical genetic algorithm (GA) [1], binary genetic algorithm (bGA) [1], differential evolution (DE) [12], classical particle swarm optimization (PSO) [16], self-adaptive neighbourhood search DE

(SaNSDE) [16] and cooperative coevolutionary particle swarm optimization (CCPSO) [8]. The values of the parameters c_1, c_2, and w used in PSO were selected based on [8]. The crossover and mutation probability are set to 0.5 [1]. The initial parameters of DE and SaNSDE were set based on [16].

Besides above parameters, there still exist common parameters such as population size, initial group size and the max iterations for all algorithms, they are shown in Table 1.

Table 1. Common parameters

Data	Pop size	Initial group size	Max iteration
5×5	100	10	1000
10×10	1000	10	5000
15×15	1000	50	5000
20×20	5000	100	10000

4.2 Results and Discussion

(1) Bayesian-based grouping mechanism
In order to prove the effectiveness of Bayesian-based grouping mechanism, we did experiments focus on fJSP and S-fJSP. And we used PSO as the basic algorithm, doing experiment on classic PSO (classical), PSO with fixed grouping (fixed), set-based grouping PSO (set) and Bayesian-based (Bayes) grouping PSO respectively. The objective is to minimize the mean value of makespan, the result are shown in Table 2. We can get the conclusion from the table that, Bayesian-based grouping mechanism has better performance no matter for fJSP or S-fJSP on mean value of makespan.

(2) Effectiveness of BNPSO
We compared BNPSO GA, bGA, PSO, DE, CCPSO and SaNSDE on the mean value and variance of makespan, the results are shown in Table 3. Firstly we can get the conclusion that, for encoding methods, real number encoding method (GA) performances better than the binary encoding method (bGA). The reason is that real number can use only one number to represent one variable, hence, for equal-length particle, real number encoding particle has larger search space. Secondly, for the basic algorithms which contains: GA, bGA, DE, PSO, we can find that GA performs better than PSO in most time, however, due to the characters of GA, co-evolution GA performs bad unfortunately. Compared to improved algorithms: CCPSO and SaNSDE, we get the conclusion that BNPSO performs better than them on most evaluation indexes.

Table 2. Results of fJSP and S-fJSP for grouping mechanism

FJSP	Target	Classical	Fixed	Set	Bayes
5×5	Mean	245	282	212	157
	Target	273.4	340.47	279.86	217.2
10×10	Mean	742	631	694	569
	Target	748.9	686.47	766.86	696.1
15×15	Mean	1209	1080	1074	979
	Target	1277.4	1174.37	1232.7	1132.63
20×20	Mean	1784	1715	1677	1372
	Target	1794.83	418.9	1774.96	1555.06
S-fJSP	Target	Classical	Fixed	Set	Bayes
5×5	Mean	273.43	261.53	279.86	217.2
	Target	831.64	768.32	708.048	774.35
10×10	Mean	748.966	730.523	766.86	696.1
	Target	468.4322	1489.65	2105.64	3148.75
15×15	Mean	1277.4	1210.4	1232.7	1132.63
	Target	1886.37	2832.33	4389.41	5596.76
20×20	Mean	1794.83	1674.83	1774.96	1555.06
	Target	216.672	1198.022	4211.4	6958.72

Table 3. Mean value and variance value of makespan for all algorithms

Scale	Target	GA	BinaryGA	DE	PSO	SaNSDE	CCPSO	BNPSO
5×5	Mean	291.5	305.76	314.1	273.43	310.66	279.86	217.2
	Variance	573.25	1068.44	318.75	831.64	203.2	708.048	774.35
10×10	Mean	778.4	834.06	809.9	748.966	805.7	766.86	696.1
	Variance	2028.1	1905.72	907.29	468.4322	1102.74	2105.64	3148.75
15×15	Mean	1301.5	1365.8	1316.06	1277.4	1238.166	1232.7	1132.63
	Variance	1242.71	3612.49	304.46	1886.37	548.595	4389.41	5596.76
20×20	Mean	1832.46	1869.266	1816.96	1794.83	1778.533	1774.96	1555.06
	Variance	2841.64	2370.02	705.09	216.672	1170.51	4211.4	6958.72

5 Conclusion

This paper presents an effective BNPSO, which solves the S-fJSP with the uncertainty of processing time. It minimized the expected average makespan within reasonable time. With the framework of the proposed BNPSO, we construct BN structure according to the data showing the relationship among variables and adjust the group scene based on the independence shown by structure. We first proved the effectiveness of BN-based grouping mechanism, and then, we compared proposed algorithm with other compared famous algorithms, the results shown that, BNSPO performed better than other algorithms. In our future work, we will extend BNPSO to adapt to real case study based on multiobjective stochastic flexible job- shop models (moS-JSP).

Acknowledgements. This work is partly supported by the National Natural Science Foundation of China under Grant 61572100 and in part by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS) No. 15K00357.

References

1. Gao J, Sun L, Gen M (2008) A hybrid genetic and variable neighborhood descent algorithm for flexible job shop scheduling problems. Comput Oper Res 35(9):2892–2907
2. Gen M, Cheng R (2000) Genetic algorithms and engineering optimization. John Wiley, NewYork
3. Gu J, Gu X, Gu M (2009) A novel parallel quantum genetic algorithm for stochastic job shop scheduling. J Math Anal Appl 355(1):63–81
4. Hao X, Chen X et al (2011) Cooperative Bayesian optimization algorithm: a novel approach to simultaneous multiple resources scheduling problem. In: Second international conference on innovations in bio-inspired computing and applications, pp 212–217
5. Lawler EL, Lenstra JK, Kan AHGR, Shmoys DB (1989) Sequencing and scheduling: algorithms and complexity. Handb Oper Res Manag Sci 4(2):445–522
6. Lei D (2010) A genetic algorithm for flexible job shop scheduling with fuzzy processing time. Int J Prod Res 48(10):2995–3013
7. Li J, Zhang JQ et al (2014) Team of Bayesian optimization algorithms to solve task assignment problems in heterogeneous computing systems. In: IEEE international conference on systems, man and cybernetics, pp 127–132
8. Li X, Yao X (2012) Cooperatively coevolving particle swarms for large scale optimization. IEEE Trans Evol Comput 16(99):1
9. Liu B, Wang L, Jin YH (2005) Hybrid particle swarm optimization for flow shop scheduling with stochastic processing time. In: Proceedings of the Computational Intelligence and Security, International Conference, CIS 2005, Xi'an, China, 15–19 December 2005, pp 630–637
10. Omidvar MN, Li X et al (2010) Cooperative co-evolution for large scale optimization through more frequent random grouping. In: Evolutionary Computation, pp 1–8
11. Potter MA, Jong KAD (1994) A cooperative coevolutionary approach to function optimization. Lecture Notes in Computer Science, vol 866, pp 249–257
12. Price K, Storn RM, Lampinen JA (2005) Differential evolution: a practical approach to global optimization. Natural computing series, vol 141(2), pp 1–24
13. Tavakkoli-Moghaddam R, Jolai F et al (2005) A hybrid method for solving stochastic job shop scheduling problems. Appl Math Comput 170(1):185–206
14. Van den Bergh F, Engelbrecht AP (2004) A cooperative approach to particle swarm optimization. IEEE Trans Evol Comput 8(3):225–239
15. Yang J, Xu H et al (2011) Task scheduling using Bayesian optimization algorithm for heterogeneous computing environments. Appl Soft Comput 11(4):3297–3310
16. Yang Z, Tang K, Yao X (2008) Self-adaptive differential evolution with neighborhood search. In: Evolutionary Computation, pp 1110–1116
17. Zhou R, Nee A, Lee H (2009) Performance of an ant colony optimisation algorithm in dynamic job shop scheduling problems. Int J Prod Res 47(11):2903–2920

Exchange Rate Movements, Political Environment and Chinese Outward FDI in Countries Along "One Belt One Road"

Wenjing Zu[1] and Haiyue Liu[2](✉)

[1] School of Economics, Zhejiang University, Hangzhou 310013, People's Republic of China
[2] Business School, Sichuan University, Chengdu 610065, People's Republic of China
seamoon@vip.sina.com

Abstract. This paper explores the role of political environment (PE) and exchange rate (ER) movements in determining China's outward foreign direct investment (OFDI) for a period of 2003–2015. It includes 65 countries along "OBOR" (One Belt One Road) to build up a panel data base. ER level, ER volatility and Political Environment are examined in relation to Chinese OFDI in these "OBOR" countries. The empirical results show that RMB appreciation is associated with higher Chinese OFDI flow but Chinese OFDI is not sensitive to ER volatility and PE instability. Considering the growing importance of the "OBOR" initiative, we believe that our research results shed light on the trend of Chinese OFDI.

Keywords: Chinese outward foreign direct investment · "One Belt One Road" Initiative · Exchange rate · Political environment

1 Introduction

Since China adopted its "going out" policy in 2001, its OFDI flows have grown rapidly, reaching more than $100 billion in 2013. As an upgraded strategy for foreign investment, the "OBOR" was initiated in 2013 to better serve China's "going out" policy. In the area defined by this "OBOR" initiative, more than 60 countries have expressed their interest in cooperating with this program. In addition, the China-led Asian Infrastructure Investment Bank (AIIB) was created to provide professional, efficient financial support for the tremendous infrastructure needs under the "OBOR" initiative. At present, a total of 57 countries covering five continents have been approved as founding members of this international financial institution. As an Asian regional multilateral development agency, AIIB will play a constructive role in strengthening economic growth engine of infrastructure construction, improving the capital utilization efficiency and raising the regional development level.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_121

However, it is becoming increasingly clear that the global market is full of uncertainties that can undermine even carefully laid plans. The risks that Chinese firms are facing when they are encouraged to invest abroad should not be overlooked. According to Wang [26], the major risks that Chinese firms deal with during outward investment are political and financial risks, which accounted more than 60% of the total risks. The risks are mainly influenced by the political environment and exchange rate movements which in turn affect the volume and regional allocation of Chinese OFDI. Therefore, this paper examines these two major risks and their influences on Chinese OFDI. The panel dataset includes 65 countries along the "OBOR". This paper uses an index that is composed of 6 weighted indicators covering both political and social attributes to measure PE risk. The ER risks are calculated by measuring ER level and volatility.

The contribution of this paper is twofold. Firstly, little paper discussed China's OFDI in "OBOR" countries, let alone analyzed the impact of financial and political risks. Unlike the previous studies which are mainly focused on the Chinese FDI flows to developed countries or emerging economies, our study looks at the countries along the "OBOR" diversified in South and Central Asia, Central and Eastern Europe and Africa. Therefore, this paper contributes to the existing empirical literature on Chinese OFDI in a new region. Secondly, our result reveals that Chinese OFDI seems not to be affected by PE and ER volatility, which is not in line with the previous literature.

The remainder of this paper is organized as follows: Sect. 2 reviews the relevant theoretical and empirical literature. Section 3 provides a descriptive overview of Chinese OFDI in the "OBOR" countries. Section 4 describes the data. Section 5 presents the empirical specification and results. Section 6 concludes and relates policy implications.

2 Literature Review

Theoretical and empirical studies have looked at the foreign investment behavior of multinationals and have identified exchange rate movements and political environment as important determinants of outward FDI. We conclude these studies as follows.

2.1 The Impact of Exchange Rate Movements on FDI

The interest in the impacts of exchange rate and its volatility on international capital flows such as FDI is growing among policy makers. As for empirical research, quite a lot of scholars try to investigate the relationship between exchange rate and FDI by taking developed countries and newly emerging countries as sample, which is critical to the formulation of FDI polices.

With regard to the exchange rate level, one strand of the literature emphasizes the positive correlation between the appreciation of source country currency and FDI outflows based on the relative wealth effect, the expectation of future profitability and capital market imperfection. Schmidt and Broll [23] showed that the real exchange rate level has a positive effect on outward FDI flows in nine industries from the US. Takagi and Shi [24] used the panel data of Japanese FDI flows to nine dynamic Asian economics during 1978–2008 and found the current depreciation of host country currencies significantly increased FDI inflows from Japan. However, another strand emphasizes the ambiguous impact of currency appreciation. Pain and Welsum [21] provided no clear conclusions as to the impact of exchange rate movements on FDI due to the types of multinational activities undertaken in different countries. Lee and Min [17] took Korea's eight major FDI source countries in three different regions in the world as samples and tried to identify the changing behavior of foreign investors in Korea following the 1997 crisis. The change in FDI in response to exchange rate level is quite mixed, which is consistent with recently developed real option-based FDI theory.

There are contradictory conclusions of exchange rate volatility influences on FDI flows in theoretical and empirical literature. The risk averse multinationals consider the information-searching cost and may decide to put off investing overseas when faced with dramatic volatility of exchange rate. Udomkerdmongkol and Morrissey et al. [25] found evidence for a negative impact of exchange rate variation on FDI of US in emerging countries. However, another group of studies highlight the positive impact of exchange rate volatility on FDI. Cushman [6], in his theoretical models, concluded that the exchange rate uncertainty may positively affect FDI. In response to risk, the multinational firm reduces exports to the foreign country but offsets this somewhat by increasing foreign capital and stimulating direct investment. Goldberg and Kolstad [11] also came to a similar conclusion, namely, an increase of the uncertainty stimulates FDI. Deseatnicov and Akiba [8] employed a panel data analysis of 56 developed and developing countries (country and industry level) and found that exchange rate volatility positively affect Japanese FDI activities for all industry. The results showed that Japanese MNCs could tolerate a slight increase in the exchange rate volatility because the level of it may be far enough what is necessary.

Although Chinese OFDI has become a more interesting topic, relatively few empirical studies have been conducted. Some research has made contribution to the literature by examining the major determinants of Chinese OFDI including ER risks as well. But the results are controversial and apparently they did not include samples from enough counties. For example, Jin [14] found that the appreciation of RMB promotes FDI after the reforms in the ER regime in 2005. Hu [13] used a panel data including 49 countries from 2003 to 2010. They found that Chinese OFDI was positively related to ER level and negatively related to ER volatility. Liu and Deseatnicov [19] found out that Chinese MNCs tend to invest in locations with higher financial uncertainty because exchange rate volatility increases the competitive advantage of Chinese MNCs in developing countries with respect to MNCs from developed countries motivating an increase in Chinese OFDI.

2.2 The Impact of Political Risk on FDI

When entering foreign market, multinationals have to adapt their optimal investment strategy to the requirements of local political environment that may give rise to varying political risks. Different political environment and its impacts on economic activity have received substantial attention in the recent literature on FDI. There are a lot of scholars trying to prove that any form of political risk negatively affects the FDI profitability and incentive of an MNE as a whole. Busse and Hefeker [3] took 83 developing countries as samples and tried to prove the negative linkages among political risk, institutions, and FDI inflows. They also found out that political risk and institutional indicators matter the most when multinational corporations confront decisions about where to invest in developing countries. Hayakawa et al. [12] used the overall FDI inflows for 89 countries to empirically investigate the effects on inward FDI of various components of political risk. The results showed that internal conflict, corruption, military in politics, and bureaucracy quality are inversely related to inward FDI flows. Cezar and Escobar [4] tested their model using inward and outward FDI data on OECD countries and found that institutional distance reduces both the likelihood that a firm will invest in a foreign country and the volume of investment it will undertake. And they concluded that firms from developed economies adapt more easily to institutional distance than firms from developing economies.

Nevertheless, some previous studies implied that better political environment may not attract more FDI inflows. Kurul [16] constructed a composite institutional quality index to identify the overall impact of institutions on FDI inflows based on 126 countries over the period 2002–2012. The empirical results revealed that institutional quality affects FDI positively only after this measure exceeds a certain threshold value. In particular, Buckley et al. [2] found that Chinese foreign investment may be led by China's political and ideological heritage and be directed to ideologically similar countries which have higher records of political risk. Yeung and Liu [27] also supported this conclusion. They hold the view that Chinese state-owned transnational firms, to a great extent, played a crucial role in foreign investment. Partly due to China's unique and changing political-economic context, many state-owned transnationals are often involved in infrastructure projects and resource extraction industries in some developing economies with poor government institutions as a means of economic diplomacy.

3 The Background of Chinese OFDI in the Countries Along the "OBOR"

3.1 The Regional Distribution of the "OBOR" Initiative Area

The "OBOR" initiative focuses on bringing together China, Central Asia, Russia and Europe (the Baltic); linking China with the Persian Gulf and the Mediterranean Sea through Central Asia and West Asia; and connecting China with Southeast Asia, South Asia and the Indian Ocean (Table 1). In its largest definition, the "OBOR" initiative involves more than 60 countries, with a total

Table 1. The geographic distribution of countries along the "OBOR"

Region	Major countries
Northeast Asia	Mongolia, Russian Federation
Central and Eastern Europe	Bulgaria, Belarus, Austria, Belgium, Czech Republic, Romania, Poland, Ukraine, Hungary
Western Asia and the Middle East	Georgia, Turkey, United Arab Emirates, Israel, Egypt, Oman, Qatar, Iraq, Iran, Syrian Arab Republic, Saudi Arabia, Yemen, Rep.
Central Asia	Uzbekistan, Turkmenistan, Tajikistan, Kazakhstan, Kyrgyzstan
Southeast Asia	Vietnam, Malaysia, Myanmar, Cambodia, Thailand, Brunei Darussalam, Indonesia, Lao PDR, Philippines, Singapore
South Asia	Pakistan, India, Jordan, Bangladesh, Sri Lanka, Afghanistan

population of 3.08 billion, accounting for about 44% of the world's population. The total GDP of this area reaches $12.8 trillion, accounting for 17% of the global economy.

3.2 Chinese OFDI Along the "OBOR"

(1) The pattern of Chinese OFDI
In recent years, Chinese OFDI along the "OBOR" countries has increased rapidly as economic ties strengthened. Figure 1 describes the Chinese direct investment stock trends in the countries involved in the "OBOR" initiative from the period of 2003 to 2015. By the end of 2015, total flows of OFDI by Chinese enterprises to the countries covered by "OBOR" initiative reached $18.93 billion, with an increase of 38.6%, accounting for 13% of the total flows.

China's OFDI varies in different economic entities. Chinese direct investment along the OBOR concentrated on the developing countries from 2003 to 2015, accounting for 75% of the total share. Among these developing countries, ASEAN countries have absorbed more than 40% of China's OFDI and have become some of China's most important economic partners. In 2015, China's OFDI flows to ASEAN countries experienced rapid growth and reached US$ 14.604 billion. By the end of 2015, there had been six ASEAN countries included in China's top ten OFDI targets along the OBOR (Table 2). ASEAN economies have been at the forefront of the emerging markets success story of the past two decades. It is obvious that ASEAN countries occupy important positions in Chinese foreign investment. Additionally, Chinese overseas investing activities in the transition region have increased by 20% over the last couple of years. Rich natural resources and the expansion of markets in this region are and will continue to be a considerable attraction for foreign multinationals. Recent growth for this transition

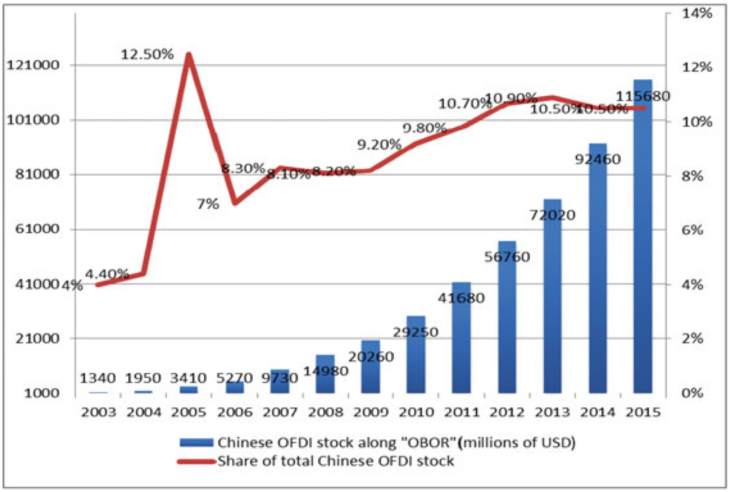


Fig. 1. Chinese OFDI Stock into the "OBOR" countries (2003–2015) Sources: Statistical Bulletin of China's Outward Foreign Direct Investment.

Table 2. Top ten countries of Chinese OFDI stock along the "OBOR" by the end of 2015

Rank	Country	Chinese OFDI stock (million dollar)	Percentage (%)
1	Singapore	31984.91	27.65
2	Russia	14019.63	12.12
3	Indonesia	8125.14	7.02
4	Kazakhstan	5095.46	4.41
5	Laos	4841.71	4.19
6	United Arab Emirates	4602.84	3.98
7	Myanmar	4258.73	3.68
8	Pakistan	4035.93	3.49
9	India	3770.47	3.26
10	Cambodia	3675.86	3.18
Total	84410.68	72.98	

[a] Sources: Statistical Bulletin of China's Outward Foreign Direct Investment (2015)

region is highest among all regional groups. Kazakhstan, Mongolia and Russia are currently among the largest FDI recipients in the region. It is worth noting that the transitional region might become a new spotlight for Chinese foreign investment, especially motivated by the "OBOR" initiative.

(2) The industry structure and investment project of Chinese OFDI
There has been a trend of diversification in the industrial structure of Chinese OFDI along the "OBOR". In 2005, the Chinese large-scale project investment

along the "OBOR" only involved energy industry that was mainly dominated by oil and supplemented by natural gas and coal. Between 2006 and 2008, China's large-scale project investment extended to the metal ore industry, real estate, transportation and other industries. The Chinese enterprises further expanded their investment to high-tech, agriculture, finance, and chemical industries from 2009 to 2015. These changes in China's OFDI along the "OBOR" show it has experienced a steady process of ascension. In general, the focused industry is energy. Metal ore, real estate and transportation rank second to fourth. Agriculture, chemical and high-tech industries account for a smaller part of total shares.

The "OBOR" initiative makes the interconnection of infrastructure as a breakthrough in cooperation. Driven by this initiative, China actively develops high-speed railway networks, expressway networks, and regional aviation networks. In 2015, China plans to invest infrastructure construction along the "OBOR" for 1.04 trillion Yuan, especially in railways, water conservancy and port engineering, and airports, in that order (Fig. 2).

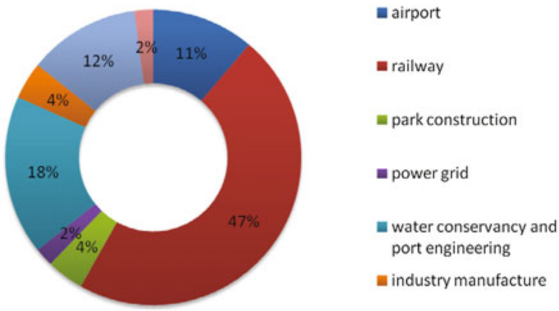


Fig. 2. China planed investment projects in "OBOR" countries (2015) (Sources: Minsheng Securities)

(3) The investment subject of Chinese OFDI

Under the "OBOR" initiative, government participation makes up the most important part while social capital participation remains a relatively small part. From the point of investment scale, central state-owned enterprises are the main force of China's OFDI along the "OBOR" while local private enterprises only play a supplementary role. By the end of the first half of 2014, the central enterprises' national large-scale projects investment stock in this "OBOR" area reached US $86.45 billion, accounting for 67.4% of the total share. Especially, the OFDI stock from state-owned enterprises belonging to SASAC was US $78.22 billion, making up 90.5% of the central enterprises' investment. In addition, local enterprises large-scale projects investment stock in this area was US $41.9 billion, accounting for 32.6% of the total share.

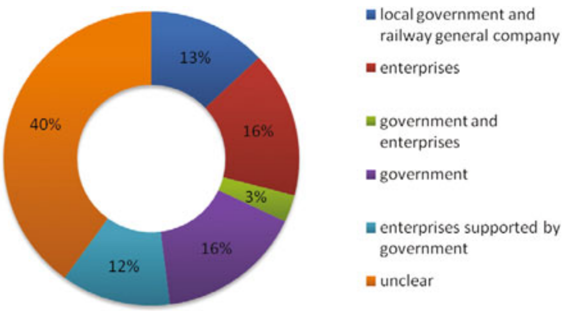


Fig. 3. Chinese Investment Subjects in "OBOR" Countries (By the End of the First Half of 2014, Sources: Minsheng Securities)

4 Data

4.1 Sample Selection

The aim of this paper is to examine the impact of exchange rate movements and political environment on Chinese OFDI along the "OBOR". Since the Ministry of Commerce, National Bureau of Statistics, and State Administration of Foreign Exchange started to publish a detailed Statistical Bulletin of China's Outward Foreign Direct Investment from 2003, our dataset contains the Chinese OFDI net flows to selected countries for a period of 2003–2015.

We realize that a large number of country samples are critical to our analysis. The Land and Maritime Silk Roads pass through over 60 countries with transmission of cultures, religion and trade. However, these vast networks cover so many developing and small countries, the detailed official FDI data of which cannot be found in Statistical Bulletin of China's Outward Foreign Direct Investment. As a result, our study met difficulties in collecting enough samples of countries along the "OBOR". So we include countries that are not geographically located on the "OBOR" but also belong to its coverage such as European countries, Africa and Far East. In this paper, the total sample consists of 65 countries on the basis of all available information (See Table 4 in Appendix).

4.2 Data Description

The dependent, independent and control variables are as follows. The summary statistics of the empirical variables is reported in Tables 5 and 6 in Appendix shows the sources of the data.

(1) Dependent variable

The dependent variable for our analysis is the outward FDI flows to the host country by Chinese multinationals during the sample period. We use LRFDI, the natural logarithm of real FDI flows from China to each recipient country expressed in ten thousand dollars (the CPI is used to deflate the nominal values, with 2010 as the base year), as a measure of FDI. To understand the details of China's OFDI, we downloaded an annual country-level data from Statistical Bulletin of China's Outward Foreign Direct Investment. This official government data repository for all Chinese firms (including Limited companies, Private enterprises, State-owned enterprises, Foreign-invested or Joint-Venture firms) and has rich information including the destination of OFDI, the various sectors, the name of the investor in China and regional origin of the investor. This is so far the most reliable information at country-level in long time-series.

(2) Independent variables
This paper mainly investigates the effect of the political environment and exchange rate movements on Chinese OFDI in the "OBOR" countries. The measurement of independent variables proceeded as follows.

(a) Exchange Rate variables
We also collect a set of observable monthly nominal exchange rates (RMB/local currency, using US dollars as the transformation instrument) from selected countries which can be obtained from the IMF International Financial Statistics database. In order to obtain the real exchange rates, we use host country's CPI index to deflate the nominal values, with 2010 as the base year. From this data, we study the level and volatility of exchange rate to explore their impacts on Chinese OFDI.

- LMEANR is the natural logarithm of the level of real exchange rates. We calculate the first moment (average) of monthly real exchange rates around year t (defined to include all monthly observations during year $t-1$ and year t) for host country.

$$\text{LMEANR} = \log\left(\sum_{\text{i}=1}^{24} {x_i}^2/24\right).$$

- LVARR is the natural logarithm of the volatility of real exchange rates. We calculate the second moment (standard deviation) of real exchange rate level around year t (defined to include all monthly observations during year $t-1$ and year t) for host country.

(b) Political Environment
Political environment, however, varies across countries and regions, which provides us an opportunity to test the impact of the severity of PE risks on outward FDI by Chinese multinationals. In this paper, we use PE to represent the evaluation of host country's political environment. As one of the most widely accepted system of completely independent political risk forecasting, the Political Risk Services Group (PRS) offers us a general risk model to translate political risk assessments into numbers and calculate risk ratings for its database of 140 developed, emerging, and frontier markets. The Political Risk Rating comprises 6 same weighted variables covering both political and social attributes in order to focus on the needs of the particular investing firm (see Appendix 1d). These

actions can result in threats or do harm to the business climate. Thus, we calculate the sum of PRS 6 indicators to capture the features of political environment in host country. An increasing value (from 0 to 6) represents lower political risk and better political environment.

(3) Control variables
We look at the macroeconomic determinants of outward FDI. The host country's CPI index is used to deflate the nominal values, with 2010 as the base year. Following previous research, we tested main factors as follows:

- LRGDP is the natural logarithm of host country's real GDP (Deflated by CPI) based on current US dollars in a given year, which reflects the host country's future market potential or absorptive capacity.
- NATURAL is the natural logarithm of total natural resources rents (% of GDP). The total natural resource rents are defined as the sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents, and forest rents. The estimates of natural resources rents are calculated as the difference between the price of a commodity and the average cost of producing it. This indicator could be used to measure the abundance of natural resource in the host country.
- LRW is the natural logarithm of real wage in host country. GDP per capita based on purchasing power parity (current US $) serves as a proxy for real wage due to the difficulties of getting average wage data for most of the "OBOR" countries. It weighs the labor cost in host country and is widely used to measure the advantage of human capital in attracting foreign manufacturing firms.
- LOPEN is the natural logarithm of the sum of annual imports and exports volume of goods and services (% of GDP). It stands for the degree of host country's openness in the FDI recipient economy, which is considered to be an important factor affecting FDI-related barriers.
- LTECH is the natural logarithm of host country's high-technology exports (% of real GDP). High-technology exports are products with high R&D intensity, such as in aerospace, computers, pharmaceuticals, scientific instruments, and electrical machinery. Data are in current U.S. dollars. This variable could measure the level of technology in host countries, which becomes the one of consideration in Chinese MNCs foreign investment activities.
- PLCY is a dummy variable that can be interpreted as the effect of the "OBOR" initiative. The PLCY dummy takes a value of zero if the time period is 2003–2013 and one otherwise. We splits the sample around 2013 because it is believed that Chinese OFDI activities in the countries along the OBOR are strongly driven by the "OBOR" initiative that was first launched by Chinese government in 2013.

5 Empirical Specification and Results

A linear dynamic panel-data model is used in this paper taking into consideration the lagged FDI effect on current FDI. We use the System GMM (SYS-GMM) estimator due to its capacity to deal with lagged values of the endogenous variables as poor instruments for first differences. With characteristics of missing observations for some countries and years, our dataset represents as an unbalanced panel data. Thus, we use "forward orthogonal deviations" variables transformation proposed by Arellano and Bond [7] in order to increase the number of observations used in the analysis.

As for instruments choosing, the independent variables were treated as strictly exogenous. So the system estimators use the first difference of all the exogenous variables as standard instruments, and the lags of the endogenous variables to generate GMM-type instruments. Finally, we added time dummies in order to increase the likelihood of no correlation across individuals in the idiosyncratic disturbances assumption to hold as suggested by Roodman [22].

As robustness check to our SYS-GMM regressions we report feasible generalized least squares (FGLS) estimation method with heteroscedastic error term. We include time dummies to the FGLS estimations as well and we exclude lagged FDI variable in order to avoid the autocorrelation problem in the FGLS method.

We estimate the equation:

$$y_{it} = \delta \times y_{it} - 1 + X'_{it}\beta + \varepsilon_{it}, \tag{1}$$

where y_{it} is the logarithm of annual outward FDI from China into a host 'country i' at time t and X'_{it} denote an (1xk) vector of exogenous variables which vary in the cross-section and in the time dimension. δ is a scalar. y_{it-1} is a lagged dependent variable. εit is a stochastic error term, which is assumed to be uncorrelated over all i and t. For FGLS estimation we omit δ y_{it-1} term.

All system GMM estimation results for Eq. (1) are shown in Table 3. In our specifications, the Hansen test of overidentifying restrictions indicates that the joint null hypothesis of valid instruments is not rejected. Besides, the FGLS results displayed in Table 3 show that the results of system GMM estimation are consistent and robust. Some interesting findings have been disclosed through the regression and we give some interpretations and evaluations for them.

5.1 RMB Appreciation Stimulates Chinese OFDI in "OBOR" Countries

ER level is found to be positive and statistically significant for the 65 selected countries. It is expected correctly signed and positively associated with Chinese OFDI. This suggests that Chinese multinationals are willing to invest abroad in these countries and areas when the Yuan appreciated with the host's currency. More specifically, ER level coefficient indicates that a 1% increase in the value of Yuan would promote an about 0.3% increase in Chinese OFDI. A considerable amount of OFDI is conducted through M&A, Yuan appreciation can systematically lower the cost of acquisitions of certain foreign assets [10] and increase the profitability of firms' foreign branches and subsidiaries [1]. Thus, an appreciation of home currency stimulates OFDI. Moreover, evidence found from the statistics suggests that there is a stable uprising trend for Yuan during the sample year.

Especially after 2005, Yuan appreciated over 30% [28]. In this case, the yuan's bias toward relatively large appreciation shocks is associated with expectation of the appreciation that leads to a lower cost of future reinvestment in the host country and stimulates the foreign investment.

5.2 Chinese OFDI Is Not Sensitive to the Exchange Rate Volatility in "OBOR" Countries

ER volatility coefficient shows a negative but insignificant relationship with OFDI, indicating that the influences of ER volatility on Chinese OFDI are ambiguous (Table 3). The possible explanation for this insignificant result is that firstly, the Chinese government is very cautious with respect to the pace of the

Table 3. Panel regression by System GMM and FGLS for the 65 "OBOR" Countries (Dependent variable: LRFDI)

VARIABLES	SYS-GMM (Two Dif.) LRFDI	SYS-GMM (Two Ort.) LRFDI	FGLS LRFDI
L.LRFDI	0.373	0.172	
	(−0.031) ***	(−0.061)**	
LMEANR	0.273	0.295	0.319
	(−0.087)**	(−0.124)*	(−0.082) ***
LVARR	−0.11	−0.15	−0.12
	(−0.061)	(−0.104)	(−0.071)
LRGDP	0.361	0.710	0.744
	(−0.092)***	(−0.172)***	(−0.07)***
NATURAL	0.012	0.038	0.045
	(−0.004)**	(−0.011)***	(−0.007)***
LRW	0.109	−0.177	−0.01
	(−0.143)	(−0.194)	(−0.105)
LOPEN	0.207	0.343	0.435
	(−0.057)***	(−0.062)***	(−0.087)***
PE	0.07	0.161	−0.346
	(−0.183)	(−0.252)	(−0.157)*
LTECH	0.039	0.03	0.116
	(−0.044)	(−0.057)	(−0.035)***
PLCY	0.352	2.074	3.720
	(−0.089)***	(−0.278)***	(−0.38)***
Constant	−7.106	−15.398	−15.454
	(−2.079)***	(−3.755)***	(−1.646)***
No. of observations	383	383	453
No. of groups	48	48	48
No. of instruments	38	38	
Group instrument ratio	1.263	1.263	
Hansen test of overrid. (P value)	0.169	0.241	

Note: Standard errors in parentheses. ***, **, * indicate that the null hypothesis is rejected at 1, 5 and 10 percent, respectively.

RMB appreciation. During our sample year, China's Central Bank intervened and managed the unexpected RMB exchange rate fluctuation. Also RMB is usually expected to appreciate gradually. Thus, Chinese multinationals pay less attention to the variance of ER than to the level of ER in managing financial risk. Secondly, unlike the multinationals in developed countries, the motivation for Chinese MNCs is not to substitute exports but to obtain higher values of assets or access natural resources in a host country. ER volatility is a sign of instability but at the same time it discourages potential investors from developed countries to involve in competition. Thus, Chinese MNCs may be indifferent to exchange rate volatility because of their motivation. Thirdly, less hedge tools and RMB investment instruments designed for safeguard the exchange risks are used in the international settlement practices for Chinese MNCs. They suffer large exchange loss due to a lack of awareness towards exchange rate volatility. This is a rather dangerous trait as China is going to make the RMB more flexible and the Chinese Central Bank will gradually quit the normal intervention.

5.3 Chinese OFDI Seems to Be Indifferent to Political Instability in "OBOR" Countries

The PE variable achieves a statistically insignificant result with a positive sign, which is not in line with previous studies from the perspective of risk premium and information asymmetries. This suggests that Chinese multinationals are not sensitive to the political instability when investing in "OBOR" countries. This result may be explained by three possible reasons. First, multinationals that are predominantly oriented toward seeking natural resource have a greater tolerance of political risk. Our empirical result shows that the abundance of fuel and ore resource in a host country has a positive impact on attracting Chinese OFDI. It indicates that Chinese OFDI in "OBOR" countries is motivated by the host's natural resources endowments. It is well known in the literature that countries rich in natural resources, on average, grow more slowly and have ineffective political institution than countries without such resources due to the "resource curse" [20]. The more Chinese OFDI a country attracts, the worse its government institution [15]. The "resource curse" effect can be confirmed by examples with extremely high political risk in our sample. There are some countries (such as Afghanistan, Iraq and Libya) in the Middle East, Central Asia and Africa, with large oil and other critical energy production. Therefore, it is inevitable that Chinese investors have to face and endure more political risk when conducting resource- seeking FDI in "OBOR" countries.

Second, in China, State-owned enterprises are the main force in foreign investment, whose investment behavior is affected by the government's willingness. Economics activities carried out by Chinese national firms are not only for their profit motive and economic efficiency, but also for distinctive political and diplomatic overtones in "economic diplomacy". Consequently, strongly supported by Chinese government, Chinese SOEs tend to be less averse to invest in countries with higher political risk compared to non-SOEs [9]. Government support may effectively compensate for the loss of the assets value and expected return due to the changes of political environment.

Third, China's political relations with potential hosts have significant influences on firms overseas investment decisions and patterns. Countries keeping better interstate relations with China tend to get more foreign investment from Chinese multinationals [18]. The "OBOR" initiative covers many countries that have developed good long-term relationships, supporting each other politically and co-operating economically. For instance, the members of SCO (Kazakhstan and Russia), some ASEAN countries and African countries have long-term strategic partnerships with China.

5.4 Natural Resources Abundance and Market Potential Attract Chinese OFDI in "OBOR" Countries

In preliminary regressions, we see that natural resource is the positive significant determinant of Chinese OFDI within the "OBOR" countries: that is, Chinese investors preferentially seek out the ownership advantage of natural resource in the "OBOR" countries. In the long term, Chinese resource-seeking foreign investment has been aimed at getting access to the scarce natural resource that China lacks to ensure the supply of domestic production and continue development. Especially, a great proportion of large acquisitions carried out by Chinese multinationals concentrates on the resource sector. Purchases of a 96.9% stake in the Russian oil company Udmurtneft by Sinopec, the equity M & A of Kazakhstan PK Petroleum Company by China National Petroleum Corporation (CNPC) and the strategic alliance between National Grid Corporation of the Philippines and State Grid Corporation of China are examples.

Moreover, according to our results, the two alternative measures of host market potential (real GDP and openness) attained significance. This finding reinforces the view that market potential and FDI outflow are associated positively [5]. Especially, new emerging economies or transition economies along the "OBOR" present more opportunities for generating profits than before. As an emerging hotspot, ASEAN is working towards regional integration aimed at one single market and free movement of goods, capital and services. It is becoming a rapidly developing destination for foreign investment. Besides, the transition economies have enjoyed the development dividend after their significant reforms of institutional and economic structures. Their increase has been well beyond that in Latin America and Africa in recent years.

5.5 A New Wave of Chinese OFDI in "OBOR" Countries Benefited from Favorable Policies

The estimated policy variable coefficient is positive and statistically significant in our specifications, which is intuitively plausible. The "OBOR" initiative proposed in 2013 is a centerpiece of China's foreign policy and domestic economic strategies, substantially giving boost to multilateral ties and economic co-operation that promote regional integration. Considering that "OBOR" initiative serves as a new engine to speed up investment facilitation, eliminate investment barriers, and push forward negotiations on bilateral investment protection agreements, it is not difficult to understand that the sizable increase of Chinese OFDI in "OBOR" countries. As the "OBOR" initiative has been put on the agenda, the spotlight falls on infrastructure construction, energy cooperation and advanced manufacturing sectors between China and the countries along the routes. For example, China has signed memorandum with 20 "OBOR" countries for institutional capacity cooperation, especially Indonesia, Kazakhstan, and Mongolia. There have been 52 projects with a total investment of $27 billion are being implemented under the "OBOR" initiative. Therefore, the increasing policy bonuses are encouraging Chinese investors to expand their presence overseas and explore new business opportunities (Table 7).

6 Conclusion and Policy Implications

The findings of our study have some policy implications for both the government and firms under the "OBOR" initiative. The empirical results reveal that ER level is a highly statistically significant determinant of Chinese OFDI while PE and ER volatility are not. Based on this result, the Chinese government should execute an active and strategy-focused diplomatic policy in the "OBOR" countries in order to better support and promote the Chinese multinationals' foreign investment. Research agencies and the Department of Commerce should cooperate to issue PE risk reference and precaution for the firms that have already invested or have the potential to invest. Industry associations should also play a role in guiding the firms to invest rationally at a steady pace.

Exchange rate risks should be taken into consideration as China is moving towards a more flexible exchange rate regime. Especially, China's Central Bank has taken further action on the RMB internationalization, by widening the RMB exchange rate's floating band within the inter-bank market from 0.5% to 1.0%. Becoming effective on April 16, 2012, this policy has been seen as a strong signal to the world that the RMB exchange rate regime is going to be driven more and more by market forces, than by government intervention. Thus, exchange rate volatility will increase as RMB exchange rate reforms are made. Financial institutions, like commercial banks, should provide more hedging tools for Chinese-outgoing-firms to reduce exchange rate risks. As for Chinese firms, it is vital to observe the ER indicators because these largely determine the safety of investment and the overall profitability calculated in home currency. In conclusion, a trend that direct investment in the "OBOR" countries will steadily increase can be foreseen. It is vital for Chinese Multinationals to be more aware of risks to conduct efficient investment abroad.

Acknowledgements. This paper is supported by Sichuan Science and Technology Department under Grant 2016ZR0017. The Usual disclaimer applies.

Appendix

Table 4. List of 65 countries used in the study

Middle East and Africa (15)	Europe (27)	East, South and Central Asia (23)
Afghanistan, Angola, United Arab Emirates, Israel, Algeria, Egypt, Liberia, Libya, Oman, Qatar, Iraq, Iran, Syrian Arab Republic, Saudi Arabia, Yemen, Rep.	Bulgaria, Belarus, Austria, Belgium, Switzerland, Czech Republic, Germany, Denmark, Spain, Finland, France, United Kingdom, Ireland, Romania, Russian Federation, Georgia, Italy, Portugal, Luxembourg, Netherlands, Norway, Poland, Ukraine, Sweden, Iceland, Greece, Hungary	Azerbaijan, Brunei Darussalam, Hong Kong, Indonesia, India, Jordan, Kazakhstan, Lao PDR, Pakistan, Philippines, Bangladesh, Sri Lanka, Singapore, Uzbekistan, Vietnam, Turkey, Thailand, Turkmenistan, Macao, Cambodia, Myanmar, Mongolia Malaysia

Table 5. Summary statistics

Variable	Obs	Mean	Std.	Dev.	Min Max
LRFDI	657	3.367761	2.630155	−4.60303	11.19835
LMEANR	729	−3.72812	2.920762	−7.48804	3.590258
LVARR	716	−6.82429	3.337829	−17.5878	3.547741
LRGDP	764	21.15035	1.648647	15.90119	24.36234
NATURAL	746	14.18228	19.19367	0.000372	92.01895
LRW	764	4.60327	1.531155	0.946498	7.054828
LOPEN	806	4.083142	1.054777	−1.03116	6.120905
PE	754	4.005506	1.051376	1.414773	5.920455
LTECH	838	4.229373	9.253167	0.000009	75.041053

Table 6. Data Sources

Variables	Sources
LRFDI	Statistical bulletin of China's outward foreign direct investment
LMEANR	International financial statistics
LVARR	International financial statistics
PE	Political risk services group
LRGDP	World Bank database NATURAL World Bank database
LRW	World Bank database
LOPEN	World Bank database
LTECH	United Nations, Comtrade Database

Table 7. Variables incorporated into Political Risk Services (PRS) index

Voice and accountability	Military in politics
Political stability and absence of violence	Democratic accountability
	Government stability
	Internal conflict
	External conflict
	Ethnic tensions
Government effectiveness	Bureaucratic quality
Regulatory quality	Investment profile
Rule of law	Law and order
Control of corruption	Corruption

References

1. Blonigen BA (1997) Firm-specific assets and the link between exchange rates and foreign direct investment. Am Econ Rev 87(3):447–465
2. Buckley PJ, Zheng P (2009) The determinants of chinese outward foreign direct investment. J Int Bus Stud 40(2):353–354
3. Busse M, Hefeker C (2007) Political risk, institutions and foreign direct investment. Eur J Polit Econ 23(2):397–415
4. Cezar R, Escobar OR (2015) Institutional distance and foreign direct investment. Rev World Econ 151(4):713–733
5. Chakrabarti R, Scholnick B (2002) Exchange rate expectations and foreign direct investment flows. Rev World Econ 138(1):1–21
6. Cushman DO (1985) Real exchange rate risk, expectation, and the level of direct investment'. Rev Econ Stat 67(2):297–308
7. Dan B (1991) Some tests of specification for panel data: Monte carlo evidence and an application to employment equations: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 58(2):277–297
8. Deseatnicov I, Akiba H (2016) Exchange rate, political environment and FDI decision. Int Econ 148:16–30
9. Duanmu JL (2011) The effect of corruption distance and market orientation on the ownership choice of MNEs: evidence from China. J Int Manage 17(2):162–174
10. Froot KA, Stein JC (1991) Exchange rates and foreign direct investment: an imperfect capital markets approach. Q J Econ 106(4):1191–1217
11. Goldberg LS, Kolstad CD (1995) Foreign direct investment, exchange rate variability and demand uncertainty. Int Econ Rev 36(4):855–873
12. Hayakawa K, Kimura F, Lee HH (2013) How does country risk matter for foreign direct investment? Dev Economies 51(1):60–78
13. Hu B (2012) RMB exchange rate and Chinese outward foreign investment-a cross-country panel analysis. Contemp Econ Res 11:77–82 (in Chinese)
14. Jin W, Zang Q (2013) Impact of change in exchange rate on foreign direct investment: evidence from China. Lingnan J Bank Finan Econ 4(1):1
15. Kolstad I, Wiig A (2009) What determines Chinese outward FDI? J World Bus 47(1):26–34

16. Kurul Z (2016) Nonlinear relationship between institutional factors and FDI flows: dynamic panel threshold analysis. Int Rev Econ Finan 48:148–160
17. Lee BS, Min BS (2011) Exchange rates and FDI strategies of multinational enterprises. Pacific-Basin Finan J 19(19):586–603
18. Li Q, Sr GL (2012) Political relations and Chinese outbound direct investment: evidence from firm-and dyadic-level tests. SSRN Electron J. doi:http://dx.doi.org/10.2139/ssrn.2169805
19. Liu HY, Deseatnicov I (2016) Exchange rate and Chinese outward FDI. Appl Econ 51:1–16
20. Mehlum H, Moene K, Torvik R (2006) Institutions and the resource curse. Econ J 116(508):1–20
21. Pain N, Welsum DV (2003) Untying the gordian knot: the multiple links between exchange rates and foreign direct investment. JCMS J Common Mark Stud 41(5):823–846
22. Roodman D (2009) A note on the theme of too many instruments. Oxford Bull Econ Stat 71(1):135–158
23. Schmidt CW, Broll U (2009) Real exchange-rate uncertainty and US foreign direct investment: an empirical analysis. Rev World Econ 145(3):513–530
24. Takagi S, Shi Z (2011) Exchange rate movements and foreign direct investment (FDI): Japanese investment in Asia, 1987–008. Jpn World Econ 23(4):265–272
25. Udomkerdmongkol M, Morrissey O, Görg H (2009) Exchange rates and outward foreign direct investment: US FDI in emerging economies. Rev Dev Econ 13(4):754–764
26. Wang Y (2015) China's evaluation of investment risks in countries along 'one belt, one road' initiative. China Opening J 181(4):15–21
27. Chung Yeung HW, Liu W (2008) Globalizing China: the rise of mainland firms in the global economy. Eurasian Geogr Econ 49(1):57–86
28. Yue LH, Qiang J, Kai TY (2016) Determination of renminbi equilibrium exchange rate, misalignment, and official intervention. Emerg Mark Finan Trade 52(2):420–433

Post-Traumatic Stress Disorder Among Survivors in Hard-Hit Areas of the Lushan Earthquake: Prevalence and Risk Factors

Zongtang Xie[1](✉) and Hongxia Xiu[2]

[1] School of Management, Northwest Minzu University, Gansu 730030, People's Republic of China
zongtangx171@126.com

[2] School of Economics, Northwest Minzu University, Gansu 730030, People's Republic of China

Abstract. To estimate the prevalence of post-traumatic stress disorder (PTSD) and assess determinants related to PTSD among survivors in areas after the 2013 Lushan earthquake in China. This study used the survey instruments included the PTSD Checklist-Civilian Version (PCL-C) and the Social Support Rating Scale (SSRS). We found that the prevalence rates of suspected post-traumatic stress disorder (PTSD) in hard-hit areas were 35.6% ($n = 500$). The incidence of PTSD was more prevalent among women (43.41%), compared with men (27.27%). Based on multivariate logistic regression analyses, female sex, Tibetan ethnicity, being injured, having family members injured and social support were significant risk factors in hard-hit areas.

Keywords: Prevalence · Risk factors · Post-traumatic stress disorder (PTSD) · Lushan earthquake

1 Introduction

On April 20, 2013 at 8:02 am, an earthquake measuring 7.0 on the Richter scale occurred in Lushan County, Sichuan Province, China. Its epicenter and focal depth were at latitude 30°308'N, longitude 102°888'E, and 14.0 km, according to the location of U.S. Geological Survey. Lushan earthquake occurred in the southwestern part of the Longmen Shan fault zone [5]. It was the strongest earthquake after the 12 May 2008, Ms 8.0 (China Earthquake Data Center) or Mw 7.9 (U.S. Geological Survey), Wenchuan earthquake in Sichuan province, China. Unlike other natural disasters, earthquakes usually occur without any warning and can cause widespread devastation and expose thousands of people to sudden bereavement, injury, loss of property, homelessness, and displacement [8]. Lushan earthquake still caused strong shaking for human society around its

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_122

epicenter. Since the source regions of the Lushan earthquake are highly populated, and with numerous mountainous, it resulted in more than 200 deaths or missing persons, more than 10,000 injuries, and huge economic losses according to the local government's official report. Besides deaths, physical injuries and economic losses, earthquakes can have serious psychological impacts for survivors as well as physical injuries and various mental health problems [1].

Many survivors of natural disasters experience post-traumatic stress disorder (PTSD) in their adjustment to the loss of resources (e.g. housing, belongings) or loved ones (e.g. family members) [1]. PTSD is the most frequently reported psychological sequelae among victims of natural disasters [15]. Numerous studies have documented the estimated the PTSD among other earthquake survivors. Through a longitudinal survey on the onset and development of DSM-IV Post-traumatic Stress Disorder (PTSD) after the Zhangbei earthquake in North China, Wang et al. reported the rate of onset of PTSD within 9 months in severely affected village (30.3%) was higher than that in lightly affected villages (19.8%) [13]. Another survey two and a half months following the earthquake in Beichuan and Langzhong Counties in Sichuan Province found that the prevalence rates of suspected PTSD in heavily (Beichuan County) and moderately (Langzhong County) damaged counties reached 45.5% and 9.4%, respectively [6].

To summarize the above studies, previous assessments among survivors of natural disasters have shown that PTSD are common, and the prevalence rates of PTSD among the hard-hit survivors were far higher than those of lightly hit survivors. Therefore, understanding PTSD is essential for identifying vulnerable populations and developing culturally specific mental health interventions [11]. For the purpose of public health emergency response, we conducted a rapid assessment of the prevalence of symptoms of PTSD and associated factors among random samples of survivors in the 5 countries of Lushan, Baoxing, Mingshan, Qionglai, and Tianquan, which were the most severely affected by the Lushan earthquake.

2 Methods

2.1 Subjects

According to the hard-hit counties list published by the Central People's Government of the People's Republic of China, we conducted a cross-sectional survey in heavily damaged counties. The hard-hit counties were Lushan, Baoxing, Mingshan, Tianquan, and Qionglai. These counties were selected because they had suffered more extensive damage than other counties in China. The inclusion criteria were as follows-having a high degree of exposure to the earthquake and experiencing the complete process of the earthquake-with a fair distribution of sex, age, and place.

The survey teams were temporarily established, consisted of well-trained master's level psychology students. They participated in a 6-day training program that included lectures describing the study protocol and instruments, role-play interviews, and mutual discussion. The survey comprised all assessments and was administered by senior staff psychiatrists and psychologists from Sichuan University's Medical School.

Before conducting the formal investigation, a pilot test was carried out in August and September 2013, with a group of randomly selected survivors participating. Minor modifications and adjustments were made according to the feedback from the pilot test. The final version of the questionnaire was used in the formal investigation. All assessment forms were translated from English to Chinese and back-translated by a bilingual team of professionals.

From December 2013 to January 2014, master's level psychology students working as research assistants approached the participants in their own homes or in temporary accommodation. To ensure privacy, interviewers and participants were encouraged to complete the questionnaires in private places. Supervision came out on a day-to-day basis throughout the survey.

2.2 Instruments

(1) PTSD Checklist-Civilian version

PTSD Checklist-Civilian version (PCL-C) is one of the most commonly used instruments for PTSD symptoms that result from disaster related psychiatric disorders [2]. The 17 items of the checklist is corresponding to each symptom in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) PDSD criteria B (re-experiencing), C (avoidance/numbing), and D (hyperarousal). With regard to the Lushan earthquake attacks, participants indicated the extent of each symptom correlated with the earthquake in a scale from 1 (not at all bothered) to 5 (extremely severely). Criteria B was re-experience of the earthquake (5 symptoms), criteria C was avoidance and emotional numbing (7 symptoms), and criteria D was increased arousal (5 symptoms). The score in each PTSD domain was simply added up, generating a total PTSD score ranging from 17 to 85. A score of 50 was found to be the optimally efficient cut-off score to predict the clinical diagnosis of PTSD [14]. The Chinese version of PCL-C has shown good diagnostic agreement and sound reliability and validity [7]. The coefficient of Cronbach's alphas for PTSD was 0.88.

(2) The Social Support Rating Scale

The Social Support Rating Scale (SSRS) evaluates subjective social supports from friends and family [12]. Numerous empirical studies have demonstrated the protective function of social support for post-disaster mental health [12]. The SSRS has been shown to have high reliability and validity on a wide range of Chinese populations. It contains 10 items, measuring three dimensions of social support: subjective support (SS, 4 items), objective support (OS, 3 items), and support-seeking behavior (SA, 3 items). The item scores of the SSRS were simply added up, generating a total support score ranging from 12 to 66 [12]. The subjective support score ranging from 8 to 32, the objective support score ranging from 1 to 22, and the support availability score ranging from 3 to 12, respectively. Higher scores indicate stronger social support. The Cronbach alpha coefficient for social support was 0.87.

2.3 Data Analysis

In the present study, statistical analyses were performed using SPSS Version 16.0 (SPSS Inc., Chicago, IL, USA). All presented data were adjusted for clustering and assigned a population-based weighting factor, based on the population size of each cluster in the final listing of all villages and settlements. Demographic data were analyzed using frequency. Chi-square test was used to compare PTSD total scores among male and female participants. One-way ANOVA was used to compare PTSD, total scores among different age groups. Multivariate logistic regression analyses were used to investigate the factors that affect PTSD.

3 Results

Characteristics of the entire study group in hard-hit areas of the Lushan earthquake were summarized. The demographic data of the 500 participants shows in Table 1. Males comprised 48.4% of the participant population compared with 51.6% for females. Of the participants, 489 (97.8%) were of Han nationality, and 11 (2.2%) were of Tibetan nationality. The average age was 38.0 years, with 44.6% of the participants in the 31–50 year age group; 71.8% of the participants were married. Only 4 participants (0.8%) had a graduate education, 144 (28.8%) had a bachelor education and 352 (70.4%) had an elementary school education or were illiterate. Among our subjects, 18.4% had a monthly income of over 2000 yuan, 38.2% had an income of 1000–2000 yuan and 38.4% earned less than 1000 yuan. With regard to the survivors' fear during the earthquake, 65.2% reported "yes". Of the survivors, 51.7% reported that their social support was high in the past year, whereas 28.6% reported low. One hundred and eight participants (21.6%) reported that they suffered from at least one physical disease, and no participants reported suffering from psychic diseases.

Percentages of PTSD status, as well as their scores, among males and females are shown in Table 2. Based on the PCL-C total scores, participants were classified as suffering from PTSD if the score was equal to or higher than 50. The overall percentage of PTSD was 35.6% ($n = 178$). The percentage of PTSD was significantly higher among females (43.4%) than among males (27.3%), as were PCL-C total scores (32.78 vs. 26.64, respectively).

Table 3 compares PTSD status among different age groups. Overall, there was a significant difference between the age groups with regard to PCL-C total scores and rates. The PCL-C total scores were significantly higher in the 41 to 50 year age group, and the prevalence rate was higher than other age group. The lowest prevalence rates and PCL-C total scores were found in the 15 to 30 year age group.

Table 4 shows the factors that affected post-traumatic stress disorder (PTSD) symptoms by multivariate logistic regression. Various demographic and loss variables were entered into the models. The prevalence rate of probable PTSD was 35.6% ($n = 178$) based on the PCL-C cut-off score of 50 (Table 4).

Results of the multivariate logistic regression analyses indicated that the prevalence of probable PTSD was significantly higher among individuals with

Table 1. Demographic data of the study population

Variables		n	%
Location	Rural	381	76.2
	Urban	119	23.8
Gender	Male	242	48.4
	Female	258	51.6
Age (years)	15–30	191	38.2
	31–40	88	17.6
	41–50	135	27
	≥ 51	86	17.2
Mean $\pm S.D.$: 38.0 ± 13.7 Monthly income	$<$ 1000 RMB	192	38.4
	1000–2000 RMB	191	38.2
	2000–3000 RMB	85	17
	$>$ 3000 RMB	32	6.4
Ethnic group	Han	489	97.8
	Tibetan	11	2.2
	Education level		
	Graduate	4	0.8
	Bachelor	144	28.8
	No degree	352	70.4
Loss[a]	No	417	83.4
	Yes	83	16.6
	Injury of body	39	7.8
	Family member injured	42	8.4
	Family member died	7	1.4

Loss was defined as death of family member, or injury to family member or self as s result of the quake.

Table 2. Comparison of PTSD status for males and females

Items	Male	Female	P-value
	n(%)	n(%)	
PTSD			
Normal ($<$ 50)	176 (72.73)	146 (56.59)	$<$ 0.01
Abnormal ($\geq$ 50)	66 (27.27)	112 (43.41)	
PCL-C (total score)	26.64 ± 3.19	32.78 ± 5.6	$<$ 0.05

females, Tibetan people, injury of body, family member injured, middle-aged (the 31 to 40 year and 41 to 50 year age group) and elderly group (the over 51 year age group) than the control group. Comparing the 15 to 30 year age group with the other age group, the 31 to 40 year age group, the 41 to 50 year age group,

Table 3. Comparison of PTSD status among different age groups

	15–30 years	31–40 years	41–50 years	51 years	P-value
	N (%)	N (%)	N (%)	N (%)	
PTSD					
Normal (< 50)	141 (73.82)	49 (55.68)	74 (54.81)	58 (67.44)	< 0.01
Abnormal (≥ 50)	50 (26.18)	39 (44.31)	61 (45.19)	28 (32.56)	
PCL-C (total score)	23.52 ± 3.98	31.25 ± 4.47	33.97 ± 4.63	26.67 ± 2.03	< 0.05

and over-51-year age group had higher PTSD symptoms and odds ratios (OR). PTSD was significantly higher among victims who sustained an injury or whose family member injured during the quake compared with noninjured survivors. OR for risk of PTSD among victims who sustained an injury or whose family member injured during the quake were 1.88 and 1.53 (compared with that of the noninjured survivors) (Table 4).

4 Discussion

4.1 Prevalence of Probable PTSD

PTSD is still common eight months after the earthquake. Compared with the prevalence of probable PTSD among hard-hit survivors soon after the earthquake, the prevalence after eight months has declined but remains significant. The elevated prevalence rates of psychological symptoms showed that PTSD is common mental health problems in the hard-hit areas after exposure to this natural disaster and remained alarmingly high.

The present study found that approximately more than one third of participants suffered from probable PTSD (35.6%) based on the PCL-C cut-off score of 50. PTSD among survivors in hard-hit areas of the Lushan earthquake were common in our subjects. Our findings indicated a steady decline in the prevalence of PTSD over time [4]. As time went by, the reduced mental health problems among survivors in hard-hit areas may be associated with relatively good living conditions and substantial social support given by the government and other aid organizations. The prevalence of PTSD in our sample were high, compared with rates of PTSD in previous studies after earthquakes. The prevalence rates of probable PTSD among the Wenchuan earthquake victims in hard-hit areas were 26.3% [15]. The prevalence rates of probable PTSD, among the survivors in the hard-hit areas of the Yushu earthquake were 33.7% [16].

4.2 Risk Factors Associated with Probable PTSD

Many studies found numerous associations between the demographics and PTSD. Among the demographic variables of interest, this study found that women were at a higher risk of suffering from PTSD than men. When exposed to

Table 4. Multivariate logistic regression analysis of probable post-traumatic stress disorder (PTSD) symptoms among the survivors of the Lushan earthquake

Variable		Total	PTSD	
			OR (95% CI)	OR (95% CI)
Overall		500	178(35.6)	
Gender	Male	242	66(27.3)	1
	Female	258	112(43.4)	1.84(1.45 – 2.34)**
Age (years)	15–30	191	50(26.2)	1
	31–40	88	39(44.3)	1.71(1.13 – 2.58)**
	41–50	135	61(45.2)	2.11(1.09 – 4.06)**
	$\geq$ 51	86	28(32.6)	1.52(1.13 – 2.05)*
Monthly income	$<$ 1000 RMB	192	89(46.4)	1
	1000–2000 RMB	191	63(33.0)	0.79(0.29-2.17)
	2000–3000 RMB	85	21(24.7)	0.61(0.36-1.05)
	3000 RMB	32	5(15.6)	0.50(0.16-1.54)
Ethnic group	Han	489	173(35.4)	1
	Tibetan	11	5(45.5)	1.87(1.45 – 2.42)**
Education level	Graduate	4	2(50)	1
	Bachelor	144	47(32.6)	0.52(0.15-1.75)
	No degree	352	130(36.9)	0.67(0.37-1.22)
Injury of body	No	461	161(34.9)	1
	Yes	39	17(43.6)	1.88(1.47 – 2.41)**
Family member injured	No	458	160(34.9)	1
	Yes	42	18(42.9)	1.53(1.13 – 2.07)*
Family member died	No	493	176(35.7)	1
	Yes	7	2(28.6)	0.84(0.27-2.06)
	Social support	500		1.06(0.87 – 1.29)*

Based on single main effect logistic regression for each variable-OR is e^{β}. Significant at 95% confidence interval does not include 1.
$*p < 0.05$, $**p < 0.01$.

earthquake trauma, females were nearly twice as likely as males to develop PTSD (OR = 1.84). This may be partly due to cultural factors, as women in traditional societies such as China tend to be more dependent, with such disasters causing their dependence object to encounter heavy loss. Initiatives should give priority to female, as they were more likely to develop mental health problems than men.

In our study, ethnicity was another significant predictor for PTSD in hard-hit areas, which agrees with the conclusions of many previous studies [3]. China has a myriad of ethnic minority groups, including the Tibetan. This study found that the odds of PTSD (OR = 1.87), in members of the Tibetan ethnic group were higher compared with the Han ethnic group. After the earthquake, the ethnic minorities should receive more support, care, and help from the government, aid organizations, and volunteers. However, a few studies were opposite to our study [10].

In present study, being injured or having family members injured was also a significant risk factor for the incidence of probable PTSD in the hard-hit areas. The findings are consistent with the conclusion of many empirical studies that the intensity of exposure to a disaster or loss is among the most robust predictive factors for mental disorders [9].

Many studies have documented the significant relationship between social support and PTSD in the aftermath of disaster [4]. Consistent with previous studies on post-traumatic psychological health, the present study confirmed that social support has a protective function [1].

4.3 Limitations

This study has some limitations. First, these participants belonged to an extremely hard-hit region. Thus, it may not be appropriate to generalize conclusions in the present study to other earthquake-affected regions. Second, self-report instrument is used, so the participants may be over reported or underreported. Third, in response to the question about history of physical or psychotic diseases, participants generally reported a history of physical disease, and few participants reported a history of psychotic disease.

5 Conclusion

Despite the limitations, to the authors' knowledge, this study has played an exploratory role in revealing the prevalence and risk factors of probable PTSD in the hard-hit areas among Lushan earthquake survivors. The findings revealed that PTSD (35.6%) are common mental health problems in hard-hit areas even eight months after the earthquake. Female sex, Tibetan ethnicity, being injured, having family members injured and social support were significant risk factors in heavily damaged areas. The findings were also one of a handful of studies on the psychological sequelae of catastrophic natural disasters among non-Western populations.

Acknowledgements. This research was supported by the National Social Science Foundation of People's Republic of China (Grant NO. 11XMZ069), Scientific and Technological Project of Gansu Province of People's Republic of China (NO. 148ZCRA012) and the Special Fund of Fundamental Scientific Research Business Expense for Higher School of Central Government (Northwest Minzu University).

References

1. Andrews B, Brewin CR et al (2007) Delayed-onset posttraumatic stress disorder: a systematic review of the evidence. Am J Psychiatry 164(9):1319–1326
2. Elhai JD, Gray MJ et al (2005) Which instruments are most commonly used to assess traumatic event exposure and posttraumatic effects?: a survey of traumatic stress professionals. J Trauma Stress 18(5):541–545
3. Galea S, Nandi A, Vlahov D (2005) The epidemiology of post-traumatic stress disorder after disasters. Epidemiol Rev 27(1):78–91
4. Hafezalkotob A, Alavi A, Makui A (2015) Government financial intervention in green and regular supply chains: multi-level game theory approach. Int J Manag Sci Eng Manag 11:167–177
5. Jia K, Zhou S et al (2014) Possibility of the independence between the 2013 Lushan Earthquake and the 2008 wenchuan earthquake on Longmen Shan Fault, Sichuan, China. Seismol Res Lett 85(1):60–67
6. Kun P, Chen X et al (2009) Prevalence of post-traumatic stress disorder in Sichuan Province, China after the 2008 Wenchuan earthquake. Public Health 123(11):703–707
7. Li H, Wang L et al (2010) Diagnostic utility of the PTSD checklist in detecting PTSD in Chinese earthquake victims. Psychol Rep 107(3):733–739
8. Mcmillen JC, North CS, Smith EM (2000) What parts of ptsd are normal: intrusion, avoidance, or arousal? data from the Northridge, California, earthquake. J Trauma Stress 13(1):57–75
9. Panda S, Modak NM, Pradhan D (2014) Corporate social responsibility, channel coordination and profit division in a two-echelon supply chain. Int J Manag Sci Eng Manag 11(1):22–33
10. Sharan P, Chaudhary G et al (1996) Preliminary report of psychiatric disorder in survivors of a severe earthquake. Am J Psychiatry 153(4):556–558
11. Van Griensven F, Chakkraband MLS et al (2006) Mental health problems among adults in tsunami-affected areas in Southern Thailand. JAMA J Am Med Assoc 296(5):537–548
12. Wang L, Zhang Y et al (2009) Symptoms of posttraumatic stress disorder among adult survivors three months after the Sichuan earthquake in China. J Trauma Stress 105(5):879–885
13. Wang X, Gao L et al (2000) Longitudinal study of earthquake-related PTSD in a randomly selected community sample in north China. Am J Psychiatry 157(8):1260–1266
14. Xu J, Song X (2011) A cross-sectional study among survivors of the 2008 Sichuan Earthquake: prevalence and risk factors of posttraumatic stress disorder. Gen Hosp Psychiatry 33(4):386–392
15. Zhang Z, Shi Z et al (2011) One year later: mental health problems among survivors in hard-hit areas of the Wenchuan earthquake. Public Health 125(5):293–300
16. Zhang Z, Wang W et al (2012) Mental health problems among the survivors in the hard-hit areas of the Yushu earthquake. PLoS ONE 7(10):e46449

How Corruption Affects Economic Growth: Perception of Religious Powers for Anti-corruption in Iraq

Marwah Abdulkareem Mahmood[1,2(✉)], Yizhuang Tian[1], and Karrar Abdulelah Azeez[1,2]

[1] Business Administration, School of Management, Harbin Institute of Technology, Harbin 150001, People's Republic of China
[2] Business administration, Collage of Economics and Administrative, University of Kufa, Najaf, Iraq
marwahabdulkareem@yahoo.com

Abstract. Iraq is one of the most corrupt countries, as since it was part of TI's statistics occupy the last positions. The paper aims to analyze and understand the perceptions of corruption in Iraq and provide a general framework for anti-corruption by social tools. The education society has selected as study sample by questionnaires randomly at the University students reached (600) student and economics data from 1979 to 2015. The multiple regression analysis, percentages and charts are used to present the results. The findings indicate that corruption is a factor behind the deterioration of the country. The regression test show negative relationship between corruption and economic growth then, religious powers. In additional, the embezzlement and bribery are more common forms in state institutions appearances. The study value suggests answers out scope of government procedures in the fight against corruption that provide social tools relate with self-behavior for corruption problems.

Keywords: Corruption · Economic growth · Religious power · Education society

1 Introduction

The most common definition of corruption is the misuse of public office for private gain [15]. It has serious consequences in reducing the economic level and development [5]. As well as a reduction of per capita income [22,24], It hurts the poor more than the rich [10], to load the economic costs of institutions and social Large [18] and thus increases the levels poverty [3,9]. Corruption is not a problem of theft only but includes many forms such as bribery that may be in payments under the table and fraud, such as falsification of contract, assault,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_123

for example, examination papers, or wasting time or selling state secrets and nepotism seeking the benefit of friends and relatives for, misappropriation of funds the public. What makes matters worse is the inability of the beneficiaries of recording information about the services received [21]. It is worth mentioning that there is corrupt behavior that considers a rate of return, including bribery of more than 11–15 times [12]. Some divides corruption into the first two types [8, 20, 21], a large corruption, which is intended to official corruption in the upper levels. Small corrupt practices which routinely incorrect practiced by the staff at the lower levels. Countries are trying, especially in developing countries to build effective strategies to fight corruption, in spite of the existence of other regulatory tools to take the fight against corruption in Iraq, such as the Integrity Commission and the Financial Inspection Office and the offices of the inspectors and the media, but discontent and still more from the Iraqi state institutions reached levels alarming. Show us that since the beginning of the Iraq enters Corruption Perception-wide and in 2003 for occupies ranked No. 133 and in spite of increased measures to combat it, but it kept growing, because occupies ranked No. 167 in 2015. As shown in Table 1. Also, the private sector is weak and the level of economic growth and income individuals wobbling for these reasons we focused on Iraq in our study. Here, the two puzzles: first, do you that the Iraqi economy negatively affected by corruption, or is it similar to the Chinese economy, which is growing with the increasing levels of corruption? Second, what motives corruption in the country under study?

In this paper, we investigate into the impact of corruption on both economic growth and capita income in Iraq. Second, provide estimates that relative

Table 1. Classification of Iraq, according to the Corruption Perceptions Index

Years	Iraq sequence	Total
2003	113	133
2004	129	145
2005	137	158
2006	160	163
2007	178	179
2008	178	180
2009	176	180
2010	175	178
2011	175	182
2012	169	174
2013	171	175
2014	170	174
2015	161	167

Source: Transparency International publications efficiency of the procedures with the anti-corruption organizations in Iraq, the most important motives common forms of corrupt behavior. Moreover, we found that corruption is one of the most important reasons impeding economic growth in the country, economic hardship and poverty of the most important factors that drive an individual to exercise corrupt behavior. As well as the current government measures are ineffective in the fight against corruption, it may be the reason the ambiguity that is characteristic of the rules and procedures governing the functioning of the organizations. [16] refers to resort to innovation may contribute may raise overall performance. We believe that the focus on the ethical aspect may be the golden key that will end the problem or at least contributes to the minimization.

This paper has four parts, the Sect. 2 is research methodology. The Sect. 3 includes data analysis and discussion. Section 4 is conclusions.

2 Research Methodology

2.1 Subscribers Properties

The study analyzes the concept of citizens for corruption in Iraq for determination the actual behavior of society. A questionnaire was distributed to a sample of university students very (600), considering that young students are the heart and the leaders of the future. Distributing the questionnaire manually on (714) students returned them (637) have neglected them (37) to identify because of their health or for other reasons. It was the response rate (89%). We made this questioner basic on [20]. In addition, as well as data obtained (37) observations about corruption and economic growth variables from the organization of integrity and the Iraqi ministry of planning from 1979 to 2015. The method was used Multiple regression analysis, percentages, tables, and diagrams in the data analysis.

2.2 The Main Hypotheses

Bryant and Javalgi [4] found corruption affects economic growth and investment in each country. Particularly the impact on investment and promise a tax on profits in developing countries. But in South Korea found that more corruption leads to more economic growth, as it works on the development of economic strength [11]. In the same way, Ajao et al. [2] explained the impact of corruption on economics growth which emphasizes that corruption is the biggest inhibitor of growth in most countries of the world. The corrupt system gives contracts to dealers who paid the highest amount of bribes, regardless of the level of efficiency, as well as a tax in their pockets and not in the state treasury [13]. We summarize business literature about corruption by three hypotheses:

Hypothesis 1 (H1): The corruption variables negatively can decrease the growth rate of a gross domestic product.

Hypothesis 2 (H2): The corruption variables negatively can decrease the growth rate of capita income.
Hypothesis 3 (H3): The religious powers negatively can decrease corruption level.

2.3 Model Development

The Business literature suggests that corruption has a deleterious effect on economic growth through two main reasons: by directly decreasing the efficiency of administrative systems And generated an atmosphere of frustration, resentment, mistrust and Corrupt behavior carries additional financial costs for society [6,7,10]. In this section, we set up a statistical model of the relationships among corruption variables, economic growth and religious powers based on the model of Gyimah-Brempong [10] and Shadabi [23] as follow:

$$g = \gamma_0 + \gamma_1 gc + \gamma_2 ec + \gamma_3 tsc + \gamma_4 grc + \mu, \tag{1}$$

where g is growth rate of gross domestic product, gc is general corruption index, es is employees corruption index, tsc is Transaction speeding corruption index, grc is Governmental Unities regulations corruption index, γ_0 is a coefficient that estimates the growth rate of national income unassociated with corruption variables, γ_1 is a coefficient that estimates the association between growth rate of national income and general corruption, γ_2 is a coefficient that estimates the association between growth rate of gross domestic product and employees corruption, γ_3 is a coefficient that estimates the association between growth rate of gross domestic product and Transaction speeding corruption, γ_4 is a coefficient that estimates the association between growth rate of gross domestic product and Governmental Unities regulations corruption, μ is a GDP growth error term.

$$\mathrm{gcapi} = \gamma_0 + \gamma_1 gc + \gamma_2 ec + \gamma_3 tsc + \gamma_4 \mathrm{grc} + \omega, \tag{2}$$

where gcapi is the growth rate of capita income, and all other variables are as defined in the equation above, ω is a growth rate of capita income error term.

$$c = \beta_0 + \beta_1 \mathrm{drp} + \beta_2 \mathrm{irp} + \sigma, \tag{3}$$

where c is corruption index, irp is development religious power, irp is individual religion power, β_0 is a coefficient that estimates the corruption unassociated with religious powers, β_1 is a coefficient that estimates the association between corruption and development religious power, β_2 is a coefficient that estimates the association between corruption and individual religious power, σ is a corruption error term.

There is evidence that corruption negatively has affected both the growth rate of real income and income distribution [10]. We expect the coefficients of γ_1, γ_2, γ_3 and γ_4 to be positive on both the growth rate of gross domestic product and capita income, while in the third model, we expect the coefficients of β_1 and β_2 to be negative on the corruption.

3 Data Analysis and Discussion

The first and second models estimate of GDP and capita income growth rates presented in Table 2. Test statistics is consistent with our hypotheses. The coefficients of γ_1, γ_2, γ_3 and γ_4 are negative as expected and are statistically significant at $\alpha < 0.1$. The adjusted explanation coefficients are 0.57 and 0.47 respectively, where f-value greater than scheduled value. This indicates that the growth rate of GDP and capita income are negatively correlated with corruption variables. These results support *Hypotheses* 1 and 2 and agreement with [1,3,9,24].

The last model estimates of corruption level presented in Table 2 panel B. The coefficients of β_1 and β_2 are negative as expected, and are statistically significant at $\alpha < 0.1$. The adjusted explanation coefficients is 0.27. This indicates that the corruption level is negatively associated with development religious power and individual religious power by -0.39% and -0.43% respectively. This result is consistent with H3 and means the religious powers have affected corruption decrease that pushes to find new tools relate to ethics and spiritual behaviors these results fit with [26]. Shadabi [23] suggested but Contrasted with Paldam [17] Who pointed out that religious variables are not enough to predict corruption.

Table 2. Significant testing results of the study models–

Panel A. The regression analysis of corruption variables							
Variable	Coefficient	Model 1			Model 2		
		Estimate	SE	(t-statistics)	Estimate	SE	(t-statistics)
Intercept	γ_0	3.18***	0.29	(11.14)	2.82***	0.29	(9.64)
General corruption	γ_1	−0.51**	0.22	(−2.29)	−0.24*	0.05	(−5.45)
Employees corruption	γ_2	−0.10*	0.07	(−1.37)	−0.14*	0.10	(−1.42)
Transaction corruption	γ_3	−0.75***	0.18	(−4.15)	−0.90***	0.004	(−3.29)
regulations corruption	γ_4	−0.18*	0.13	(−1.35)	−0.38**	0.22	(−1.70)
R^2		0.62			0.53		
Adjusted R^2		0.57			0.47		
F-value		12.93			8.95		
Significant level		0.000			0.000		
N		37			37		
Panel B. The regression analysis of religious powers - model 3							
Intercept	β_0	3.89***	0.14	(27.43)			
Development power	β_1	−0.39***	0.04	(−11.61)			
Individual power	β_2	−0.43***	0.03	(−12.46)			
R^2		0.27					
Adjusted R^2		0.27					
F-value		112.42					
Significant level		0.000					
N		600					

Note: $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ standard error in parentheses.

In Fig. 1, We noted the higher percentage of the sample about (90%) of their responses agree with the concept that argues corruption is a major problem in the country by selected always scale about (50%) and often scale about (40%). This is a very high percentage confirms the extent of realization of the corruption Appearances in the society.

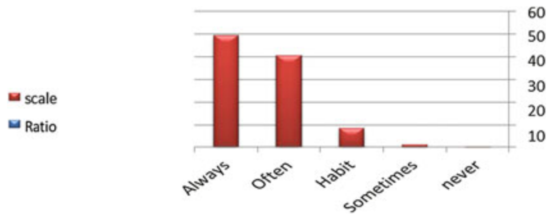


Fig. 1. The basic problem of corruption in Iraq

In comparison the impact of Corruption on economic growth. Figure 2 presents that (82%) of the participants stressed that corruption alters the level of the country and its development by selected always scale (50%) and often scale (32%). We prepared some questions to determine the level of corruption in country and to make evidence whether it contributes to change the country economics. The analytic result reports that most of the participants agreed that corruption and contributes significantly to the deterioration of the country. Lisciandra and Millemaci [14] provided evidence that corruption has negative impact on economic growth in Italian region and the relation between them is nonlinear when economic growth fall, corruption increase.

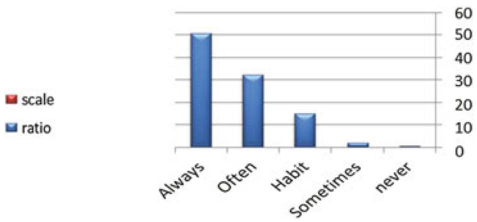


Fig. 2. The Corruption and country economics

In this study, Fig. 3 presents that the embezzlement is more famous common form to corruption phenomenon in Iraq by determined about (41.3%) of the sample. While the second form of corruption is bribery by 29.8% of their pointed, (16%) considered the third forms is favoritism, (8%) selected the extortion, and the last percentage about (4.7%) considered the information misuse is type of corruption there. Ajao et al. [2] confided that corruption forms of multiple and different gradient and conflict of interest, embezzlement, extortion, bribery and many others, which is a transfer or grab valuables and property to someone

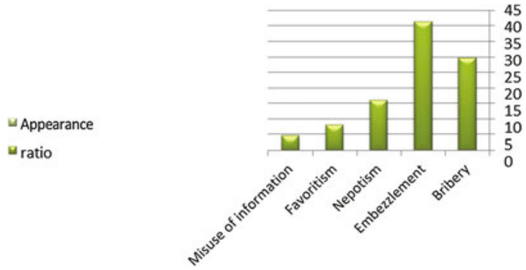


Fig. 3. The most common forms of corruption

who has no right to corruption, but used to do so by virtue of his position, and resorting to the use of misleading and erroneous information [21].

Wedeman [27] showed that widespread corruption at all levels and in the center of the large levels. This case is in our study sample consistent with studied confirmed the higher levels in the state were the most corrupt. Tabish and Jha [25] used anti-corruption strategies in Indian public sector, they are four latent constricts leadership, rules & regulations, training, and fear of punishment. The finding showed all of them are highlighted to help institutions for understanding the role of anti-corruption. Figure 4 describes the reasons of corruption based on education society. A large number of participants of the sample about) 30.5% (believes that the bribes paid based on the employee's request. Whereas about (25.2%) underlines the lack of receipt of services in the event of non-payment of a bribe, (17%) push for speeding up transactions, and (16.1%) in order to avoid problems, and finally (11.2%) pay because the treatment is not a fundamentalist.

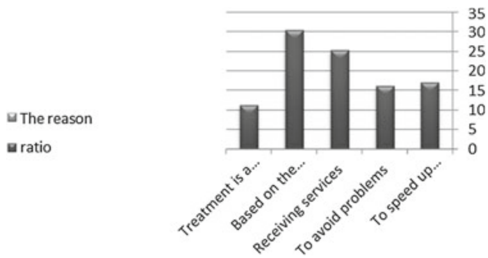


Fig. 4. Reasons for paying bribes to society situation

This study tried to know the rules and capabilities of government for anti-corruption by education society with five parts because that will make overview about the problem size and corruption effect. In Fig. 5, the higher percentage is confirmed that governmental anti-corruption measures have a low efficiency by (44.2) of the respondents, (31.7%) confirms that the measures originally inefficient and ineffective, (17.5%) considered the effect is mild, (5.2%) agreed with

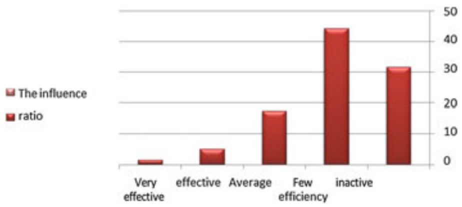


Fig. 5. The effectiveness of the procedures currently approved by government

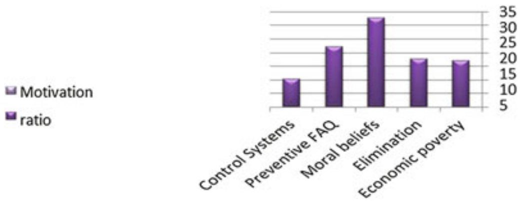


Fig. 6. The main motivations for corruption in Iraqi society

efficient and effective, and (1.3%) were a very effective and influential. Sampford, et al. [21] Delays in the implementation of judicial proceedings on the corrupt or non-implementation makes the work of that anti-corruption a few effectiveness. Lack of public confidence in the judiciary and the court and the police not to push them to give information about corruption, because they believed that the judge deal bribery or staff in court and this may be the result of the individual's personal experience in this regard. Corruption increases the uncertainty and skepticism in the judicial and legal system, the presence of authority has indisputable sovereignty opportunistic supports corruption. Also, the performance of organizations will decrease high levels of corruption, because corruption gives individuals a competitive advantage unfairly [19]. The reasons may include the failure of recent technology advances being on lower pollution emission intensity [28] (Fig. 6).

4 Conclusions

This analytic investigating aimed to expand the understanding of the impact of corruption on economic growth. Iraqi education society considers cultural focusing that has simulation key for anti-corruption. The major problem and barely missing from the country is Corruption. It also is inhibiting the development of the nation and the inhibitor of economic development and growth. Poverty or economic hardship of the individual may be the most important reasons driving for corruption for the average employee, which paid to bribe demands, on the other hand, the high cost of luxury, greed, panting behind excellence are paying the highest levels in the country to engage in corruption. We found evidence that the corruption levels are detrimental to long-run growth. This evidence is robust and persistent throughout different studies. The impact of corruption is negative with respect to economic growth at a high level for each Iraqi society.

FAQ mysterious and government actions ineffective staff gives a chance for corruption and thus the clear wording of government and striving may prevent the penetration of corruption in the institutions as well as provide the employee with a copy of the instructions when receiving the functional tasks. Accordingly, we can suggest a few considerations. A possible interpretation could be that anti-corruption by adoption non-financial measures such as moral beliefs, attitudes of safety and reliability. Finally, Corruption is not only economic but also far back in history far back in morality and must aim to change the functional ethics prevailing, and the laying of a plan put moral values purposeful organizations to clean up corruption and make it fairer, and cannot omission on the role of education society.

Acknowledgements. The authors would like to thank Kufa University in Iraq for supporting a research about data collection.

References

1. Aidt TS (2009) Corruption, institutions, and economic development. Oxf Rev Econ Policy 25(2):271–291
2. Ajao OS, Samuel D, Samuel O (2013) Application of forensic accounting technique in effective investigation and detection of embezzlement to combat corruption in Nigeria. Unique J Bus Manage Res 1:65–70
3. Beekman G, Bulte EH, Nillesen EEM (2013) Corruption and economic activity: micro level evidence from rural Liberia. Eur J Polit Econ 30(30):70–79
4. Bryant CE, Javalgi RG (2016) Global economic integration in developing countries: the role of corruption and human capital investment. J Bus Ethics 136(3):1–14
5. Chen J, Jie X, Han S (2017) Tourism economic effect divergence analysis-panel data analysis of Hunan Province. Springer, Singapore
6. Daskalopoulou I (2016) Rent seeking or corruption? an analysis of income satisfaction and perceptions of institutions in Greece. Soc Sci J 53(4):477–485
7. Gould DJ, Reyes JAA (1983) The effects of corruption on administrative performance. World Bank Staff Working Pap 580:2514
8. Graycar A, Monaghan O (2015) Rich country corruption. Int J Publ Adm 38(8): 1–10
9. Gupta S, Davoodi H, Alonso-Terme R (2002) Does corruption affect income inequality and poverty? Econ Gov 3(1):23–45
10. Gyimah-Brempong K (2002) Corruption, economic growth, and income inequality in Africa. Econ Gov 3(3):183–209
11. Huang CJ (2016) Is corruption bad for economic growth? evidence from Asia-Pacific countries. North Am J Econ Finance 35:247–256
12. Jain PK, Kuvvet E, Pagano MS (2016) Corruption's impact on foreign portfolio investment. Int Bus Rev 26:25–35
13. Kunieda T, Okada K, Shibata A (2014) Corruption, capital account liberalization, and economic growth: theory and evidence. Int Econ 139(139):80–108
14. Lisciandra M, Millemaci E (2016) The economic effect of corruption in Italy: a regional panel analysis. Regional Studies, pp 1–12

15. Lučić D, Radišić M, Dobromirov D (2016) Causality between corruption and the level of GDP. Econ Res-Ekonomska Istraživanja 29(1):360–379
16. Nazim M, Saeed R et al. (2017) Parametric analysis of leadership styles on organizational performance and the mediating role of organizational innovativeness. Springer, Singapore
17. Paldam M (2001) Corruption and religion adding to the economic model. Kyklos 54(2–3):383–413
18. U. S. S. i. S. t. U. t. o. Peace (2013) Governance, corruption, and conflict. In: A study guide series on PEACE And conflict for independent learners and classroom instructors, pp 1–54
19. Rocca ML, Cambrea DR, Cariola A (2017) The role of corruption in shaping the value of holding cash. Finance Res Lett 20:104–108
20. Sahu SK, Gahlot R (2014) Perception about corruption in public servicies: a case of BRICS countries. J Soc Sci Policy Implic 2:109–124
21. Sampford CJ, Shacklock AH, Connors C (2006) Measuring corruption. Ashgate, Farnham
22. Sanchez JI, Gomez C, Wated G (2008) A value-based framework for understanding managerial tolerance of bribery in Latin America. J Bus Ethics 83(2):341–352
23. Shadabi L (2013) The impact of religion on corruption. J Bus Inq 12:102–117
24. Stevens A (2016) Configurations of corruption: a cross-national qualitative comparative analysis of levels of perceived corruption. Int J Comp Sociol 57(4):183–206
25. Tabish S, Jha KN (2012) The impact of anti-corruption strategies on corruption free performance in public construction projects. Constr Manage Econ 30(1):21–35
26. Treisman D (2000) The causes of corruption: a cross-national study. J Public Econ 76(3):399–457
27. Wedeman A (2005) Anticorruption campaigns and the intensification of corruption in China. J Contemp China 14(42):93–116
28. Yang A, Lan X et al. (2017) An empirical study on the prisoners' dilemma of management decision using big data. Springer, Singapore

Procurement Risk Mitigation for Rebar Using Commodity Futures

Jian Ni, Wei Zhou(✉), and Dan Yang

School of Financial, Southwestern University of Finance and Economics, Chengdu 610074, People's Republic of China
zhouwei@swufe.edu.cn

Abstract. This paper presents a model to mitigate the risk of budget overruns during procurement of raw materials. A case example in rebar procurement fora typical metallurgical manufacturing company in China is used. The demand data was supplied by the company and the rebar price data was obtained from Shanghai Futures Exchanges. The novelty of this study lies in the incorporation of budgetary control into the objective function of the model. A multistage rebalancing strategy of the futures position is used to reduce the impact of unexpected changes in rebar supply price and finished product demand. The complex objective function is approximated by a quadratic expression, and suboptimal solutions are obtained for cases where raw the material price is independent of the customer demand. When compared with the company's spot buying approach without hedging, the results from the model show the proposed multistage futures balancing model can help reduce the budget overspending by as much as 32%. Further experiments are carried out to study the effects of changing price and demand volatilities, and the results confirm the usefulness of model in mitigating the underlying risk.

Keywords: Raw material procurement · Risk management · Budged control · Hedging with futures

1 Introduction and Background Literature

This paper is based on a case study in rebar procurement for a Chinese metallurgical machinery company that supplies iron and steel mills in China with drill pipe, drill head, steel making and refining equipment, and blast furnaces. The commodities markets in China are now mature and well established with a plentiful supply and fast delivery, and spot procurement is now widely used by many companies to try and avoid unnecessary inventory built-up. In recent years, however, the spot prices of steel materials have become much more volatile, and companies, if they want to remain competitive, cannot safeguard against such price volatility by raising the final product price or by imposing a contractual

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_124

surcharge on its customers. Added to this uncertainty of procurement price is the perennial problem of uncertain rebar requirements, resulting form uncertain customer demand for the finished products, especially in the prevailing worldwide economic uncertainties. This problem faced by the company in this study is not uncommon; it is faced by all manufacturers whose material procurement is exposed to severe risk arising from fluctuations in both material prices and demand for the final products.

In response to increasing raw materials price volatility caused by speculation, and to a lesser extent by volatility in material demand, many manufacturing companies are now looking at the use of financial hedging policies. Financial hedging is, of course, not an entirely new concept; much historical discussion can be found in Myers and Thompson [10], Mathews and Holthausen [8], Lence [7], McMillan [9], and Alizadeh, et al. [1]. However as suggested by Doege, et al. [5], financial hedging strategies developed for use in the world of finance are not directly applicable to risk mitigation in a manufacturing environment. With this in mind, this study is aimed at developing a dynamic multistage approach that can be used for budgetary control in material procurement under the specific business requirements of a manufacturing company. Starting with the futures contract as the instrument, the multistage hedging can be divided into two steps. First, a budget is set for a single procurement. Second, the multistage hedging is initialized and rebalanced at the end of several interim stages. The rebalancing makes use of updated information on material price and demand. With this type of dynamic rebalancing of the futures position, the financial hedge can respond fairly quickly to any unexpected changes in price and demand. At the end of the final stage, physical delivery takes place and the futures position is settled.

Another risk mitigation approach in material procurement, as reported in several studies, is the use of flexible supply contracts, such as back-up agreements [6], option contracts for physical delivery [19], and capacity reservation contracts [16,17]. It is also possible to use the flexible contracts in a synergic manner (e.g. by building portfolios) to achieve higher performance in procurement risk mitigation [18]. Generally, flexible supply contracts can alleviate the risk, but such contracts are difficult to price and may require the payment of a premium to the spot price. However, there are many situations where flexible supply contracts are not available, as is the case for the company in this study. In fact, Chod, et al. [3] employed an exponential utility to explore the relationship between operational flexibility and financial hedging opportunity. Using a similar exponential utility, Ni, et al. [12] then revealed that there can be a weak separation between the operations decision and financial hedging strategy of a firm. Other related studies on the interaction between financial hedging and operations management can be found in Okyay, et al. [20], Sayin, et al. [14], and Zhao and Huchzermeier [21].

It is only recently that researchers began to study the use of financial hedging as an approach to mitigate risk in procurement for manufacturing. In this context, Xu [20] obtained economical hedging strategies by developing an integrated framework model for procurement and inventory risk management using a financial hedge using the London Metal Exchange (LME) copper futures contracts. Another analytical framework was developed by Caldentey and Haugh [2] to investigate the appropriate hedging strategies for hedging against a possible drop in profit of a company when profit is correlated with certain returns in the financial markets. These two studies suggest that a financial hedging strategy, if carefully designed, can effectively mitigate the procurement risk. This was followed by Ni, et al. [11], who develop a multistage hedging strategy to mitigate procurement risk using a dynamic price and demand information updating process, and concluded that using appropriate information to update the hedging strategy can substantially improve the effectiveness of the hedge. Such updating makes an interim multistage rebalancing of the futures position possible, thereby providing a timely adjustment using up-to-date information. This study differs from that of Ni, et al. [11] in that it develops a hedging strategy that focuses on budgetary control. The management of the machinery company used in this study are very concerned with budgetary control and they emphasized the need to mitigate the risk of budget overruns, especially during the current economic uncertainties. With this in mind, a procurement budgetary control model (PBCM) is developed that aims at controlling the risk of excessive over budget. The PCBM model developed does not seem amenable to a direct solution, and therefore a heuristic approach had to be used to obtain efficient suboptimal solutions. Using the heuristic procedure, numerical experiments are conducted to determine the effectiveness of the multistage hedging strategy in controlling procurement spending.

In summary, this paper develops a financial hedging strategy for procurement risk mitigation in the form of an interim multistage rebalancing and dynamic information updating process, with the general objective to address the need to maintain adequate budgetary control.

2 The Dynamic Futures Rebalancing Model

2.1 Budget Control and Financial Hedging

Since the global financial turmoil that started towards the middle 2008, the management of the machinery company has been under increasing pressure to control the procurement risk arising from both volatile prices and uncertain demand. As considerable sums are being spent in procuring rebar steel required for production, the company management is also under pressure to maintain good budgetary control, and strives as much as possible to keep procurement spending in line with predetermined budgets. However, budgets can often be severely exceeded if both steel prices and product demand increase unexpectedly. When this occurs, the company may resort to banks loans to finance the overspending, but the availability of loans at reasonable interest rates is limited, mainly as a result of the current tight regulatory supervision of banks and higher Basel capital adequacy requirements. This is partly due to the economic slowdown in China over the past year or so, and perhaps also to the general downgrading of creditworthiness of private manufacturing companies as a result of this slowdown. In light of the scarcity of loans at reasonable or even unreasonable rates, the company and its peers in China need to maintain tight financial control over their procurement spending so as to reduce the type of risk that may lead to bankruptcy.

One existing approach to control procurement spending of a company is to sign long-term supply contracts with their suppliers and so to lock up the procurement price. However, this type of contract is restricted to fixed quantity deliveries, and limits the company's ability to quickly adjust to changing customer demand and will generally involve paying a premium to the spot price. China's commodity market is now quite mature and sophisticated, and commodity futures contracts are available on Shanghai Futures Exchange (SHFE) that can be used to hedge against the financial risk and to control procurement spending. As rebar is one of the raw materials that the company uses in very large quantities, the management is prepared to allow the authors to explore the use of rebar futures contracts in controlling the procurement risk that it faces. This case study is not unique, and highlights a general problem associated with commodity procurement risk management, viz., easy access to the commodity markets has encouraged many manufacturers to gradually shift to spot procurement to avoid excessive inventory built-up and achieve high flexibility in supply management. But excessive world-wide speculative activity in modern commodity markets has at the same time significantly increased price volatility. This is evidenced by the fact that only a small proportion (some 3% for copper) of futures contracts are actually settled through physical delivery, the rest can only be regarded as speculative and/or risk hedging and are settled in cash.

2.2 Risk Mitigation Model Development

The design of financial hedging strategies for commodity procurement is distinct in nature from their counterparts in the financial industry. It is not only the uncertainties in the underlying volume that arises from the changes or updates of customer orders that can extensively complicate the problem, but also the different hedging objectives (budgetary control in this paper) which require tailor-made hedging strategies to be developed. In addition, the transaction cost is usually negligible in hedging against the risk in procurement for a manufacturer [19], while this cost may be highly significant for hedging in the financial industry. This low transaction cost is mainly due to the fact that transaction costs involved in the futures contracts are very low (less than 0.02% in SHFE) in comparison to the value of the trade.

In the context of our model, the company that has to procure a certain commodity for future production (rebar for the company used in this study), where both commodity price and demand are expected to fluctuate. A multistage hedging approach is used with the aim of achieving a good overall hedge that can effectively reduce the possibility of budget overruns in procurement. Such a multistage hedging procedure is illustrated in Fig. 1. The time period or planning horizon in which we want to implement the financial hedge is T, which is divided into T stages $t = 0, 1, \cdots, T-1, T$. For the general class of problems that we are

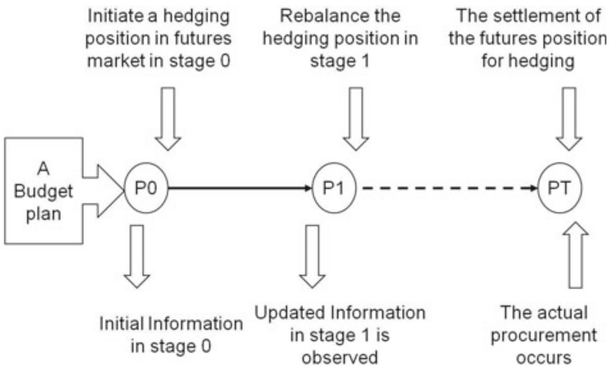


Fig. 1. The multistage hedging procedure

dealing with one stage can be a period of one week to one month, depending on the availability of updated information for both commodity price and customer demand. As shown in Fig. 1, at the beginning a budget plan is produced by the management that serves as a benchmark that the subsequent financial hedging strategy aims to meet. Using the price and demand information at stage 0, the multistage financial hedge is initialised by entering into a long position of the commodity futures that expire at the end of the final stage. At the subsequent stages, the futures position will be rebalanced to achieve the best hedge using the latest information available at each stage. At the end of the final stage, the actual procurement of the physical commodity takes place and the futures position is settled in cash. The cash settlement will completely or partially offset any overspending incurred because of the variability in both commodity price and customer demand.

The multistage hedging strategy is aimed at balancing the uncertainty of the unhedged procurement spending, C, which is defined as the product of the volatile commodity price P and the uncertain procurement quantity X, or $C = PX$. The hedge can generate a payoff from the cash settlement of the futures position, which leads to the hedged procurement spending as follows:

$$\bar{C} = C - H. \tag{1}$$

For the multistage hedging strategy, the payoff can be calculated by analyzing the stage-by-stage initialisation/rebalancing volumes y_t and the futures position z_t at each stage t $(1 \leq t \leq T-1)$. The variables y_t and z_t are related as follows:

$$z_t = \begin{cases} y_t, & t = 0 \\ z_{t-1} + y_t, & \forall\ 1 \leq t < T. \end{cases} \tag{2}$$

Given the price of the commodity futures $F_t (1 \leq t \leq T-1)$ at the beginning of stage t that will mature at the end of stage $T-1$, the payoff H can be calculated as follows:

$$H = P \times z_{T-1} - \sum_{t=0}^{T-1} F_t \times y_t. \tag{3}$$

Note that we do not consider the transaction cost. This is because of the low rebalancing frequency (weekly at most) and the low transaction cost relative to the settlement amount ratio (less than 0.02% in SHFE). From Eqs. 1 and 3, the hedged procurement spending will be:

$$\bar{C} = P\left(X - z_{T-1}\right) + \sum_{t=0}^{T-1} F_t \times y_t. \tag{4}$$

A hedging strategy is said to be effective when the hedged procurement expense $\bar{C}$ is less exposed to the volatilities of the commodity price and demand than the unhedged spending C. The design of an effective multistage hedging strategy needs to incorporate information updating processes for the material price and for the demand. These processes are used in deciding the initialisation/rebalancing volumes (y_t). In this study, the information updating process is modelled by filtration [2], a method often used to adequately model the timely availability of information. Generally, a filtration is an increasing family $\{\mathcal{F}_t\}_{t=0}^{T-1}$ of σ-algebra $\mathcal{F}_t$ [13]:

$$0 \le s < t < T \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}. \tag{5}$$

These σ-algebras $\{\mathcal{F}_t\}_{t=0}^{T-1}$ can be regarded as the available information set at each stage. Let $M_t\left(\Omega,\ \mathcal{F}_t\right)$ be the function space that is composed of all the $\mathcal{F}_t$-measurable random variables, i.e.,

$$M_t\left(\Omega,\ \mathcal{F}_t\right) = \left\{\xi \,|\xi \text{ is a random variable on } \Omega,\ \text{and } \mathcal{F}_t - \text{measurable}\right\}.$$

The hedging decisions are based on the available information at each stage, and any admissible adjusting volume y_t must satisfy the following constraint:

$$y_t \in M_t\left(\Omega,\ \mathcal{F}_t\right).$$

The multistage hedging strategy formulated below aims at keeping the procurement spending in line with an allocated budgeted. To achieve this aim, a procurement budgetary control model (PBCM) is developed which attempts minimise the expected budget overrun (EBO). The model can be stated as follows:

$$\min\nolimits_Y \text{ EBO} = E\left[\left|\bar{C} - K\right|^+\right]. \tag{6}$$

$$\text{s. t. } \bar{C} = P\left(X - z_{T-1}\right) + \sum_{t=0}^{T-1} F_t \cdot y_t, \tag{7}$$

$$z_t = z_{t-1} + y_t, \forall\ 1 \le t < T, \tag{8}$$

$$z_0 = y_0, \tag{9}$$

$$y_t \in M_t\left(\Omega, \mathcal{F}_t\right), \tag{10}$$

here the objective function $E\left[\left|\bar{C} - K\right|^+\right]$ is the EBO, which measures the average overspending with respect to the given budget K.

3 The Multistage Financial Hedging Strategy

3.1 Heuristic Solution

The discontinuity of $E\left[\left|\bar{C}-K\right|^{+}\right]$ in the objective function of the PBCM model developed above makes it difficult, if not impossible, to find a close-form solution that can be easily implemented. This inherent complexity in the PBCM model requires the use of some heuristic to generate suboptimal, though still effective and potentially useful, solutions. In this section, a heuristic is developed by first replacing the PBCM model by an α-PBCM model, where the value of α is set heuristically. The approximate model, α-PBCM, can then be solved using the closed-form results of Ni, et al. [11]; an exact solution to the approximate model. It is decided to approximate the objective function as:

$$E\left[\left|\bar{C}-K\right|^{+}\right] \cong \frac{1}{2}\left(E\left[\bar{C}\right]+\alpha\cdot E\left[\left(\bar{C}-K\right)^{2}\right]-K\right), \tag{11}$$

where

$$\alpha=\frac{E\left[\left|\bar{C}-K\right|\right]}{E\left[\left|\bar{C}-K\right|^{2}\right]}. \tag{12}$$

α needs not to be constant. However, it is expected that α will be relatively insensitive to changes in the hedging strategy used, suggesting that it should be possible to use a constant value for α in Eq. (11) and still produce a good quadratic approximation for $E\left[\left|\bar{C}-K\right|^{+}\right]$. In this study, α is determined heuristically using the following expression:

$$\hat{\alpha}=\frac{E\left[\left|\bar{C}(\hat{Y})-K\right|\right]}{E\left[\left|\bar{C}(\hat{Y})-K\right|^{2}\right]}, \tag{13}$$

where $\hat{Y}=\{\hat{y}_i\}_{i=0}^{T}$, $\hat{y}_0=E_0(X)$, and $\forall\ 1\leq t\leq T$, $\hat{y}_t=0$.

In the heuristic used, the value of $\hat{\alpha}$ is calculated on the assumption of a static hedge. It is in fact a simplified way to take into account the effect of financial hedging on the value of α. The effectiveness of $\hat{\alpha}$ is examined in Sect. 4.

When α is constant, the objective of the PBCM model can be approximated by an α-PBCM model as follows:

$$\min_{Y}\ E\left[\bar{C}+\alpha\left(\bar{C}-K\right)^{2}\right], \tag{14}$$

where Y is constrained by Eqs. (7), (8), (9), and (10).

In general, and also in the company used in this study, customer demand and material price are independent. Under this assumption the α-PBCM model can be further simplified as:

Assumption: if for every stage $t(0 \leq t < T)$ X and F_{t+1} are independent, then:

$$E_t\left[F_{t+1}X\right] = E_t\left[F_{t+1}\right] E_t\left[X\right], \forall\ 0 \leq t < T. \tag{15}$$

Under this Assumption, a close-form solution of the α-PBCM model can be obtained using the analytical results of Ni, et al. [11], because α-PBCM model has a similar quadratic objective function. While the α-PBCM and the model of Ni, et al. [11] do share a group of constraints in the multistage hedge, some modifications to two auxiliary variables are still required as follows:

$$\hat{X}_t = \begin{cases} E_{t+1}\left[X\right] - z_{t-1}, & 0 < t < T \\ E_1\left[X\right], & t = 0, \end{cases} \tag{16}$$

$$\hat{K}_s = \begin{cases} K - \sum\limits_{t=0}^{s-1} F_t \cdot y_t, & 0 < s < T \\ K, & s = 0. \end{cases} \tag{17}$$

Using the analytical results (specifically Theorem 3) in Ni, et al. [11], a closed-form solution to the α-PBCM can be obtained as:

$$y_t^* = -\frac{E\left[(F_t - F_{t+1})(F_{t+1}\hat{X}_t - \hat{K}_t + \frac{1}{2\alpha})\,|\mathcal{F}_t\right]}{E\left[(F_t - F_{t+1})^2\,|\mathcal{F}_t\right]}. \tag{18}$$

3.2 Information Updating of Price and Demand

The proposed multistage hedging strategy reacts to updates in the material price and the demand. The dynamic initialisation/rebalancing of the futures position from stage 0 to stage $T-1$ requires the predictions of the future material prices and demand levels, as well as some measure of their volatility. Such predictions are made at each stage of the planning horizon using the updating stochastic processes.

For the stochastic material price model, the long term/short term model of Schwartz and Smith [15] is employed to simulate the dynamics. When is the commodity spot price at time t, the process can be stated as:

$$\ln\left(S_t\right) = \chi_t + \xi_t, \tag{19}$$

$$d\chi_t = -\kappa\chi_t dt + \sigma_\chi dW_t, \tag{20}$$

$$d\xi_t = \mu_\xi dt + \sigma_\xi dB_t, \tag{21}$$

where dW_t and dB_t are correlated as $dW_t dB_t = \rho dt$. Equations (19), (20) and (21) produce the estimates, see Ni, et al. [11], as follows:

$$E_t\left(F_{t+1}\right) = F_t A(t), \tag{22}$$

$$E_t\left(F_{t+1}^2\right) = {F_t}^2 B(t), \tag{23}$$

where $A(t) = \exp\left\{\int_t^{t+1}\left(\lambda_\xi + \frac{\lambda_\chi}{\kappa}e^{-\kappa(T-s)}\right)ds\right\}$

$$B(t) = \exp\left\{2\int_t^{t+1}\left(\lambda_\xi + \frac{\lambda_\chi}{\kappa}e^{-\kappa(T-s)} + k(T-s)\right)ds\right\}. \tag{24}$$

In connection with the stochastic product demand, Choi, et al. [4] describe a Bayesian information update used in a two-stage application. However, for the machinery company used in this study, the Bayesian method needs to be extended to a multistage process. The raw material (rebar) required in this company for physical procurement is an aggregation of customer orders over τ stages. It is not unreasonable to assume that customer orders arrive randomly as a Poisson process. If λ is the average number of customer orders that are received in one stage, it follows that the number of orders $\tilde{N}$ arriving in τ stages is a random number from the Poisson distribution:

$$P\{\tilde{N} = k\} = (\lambda\tau)^k \frac{e^{-\lambda\tau}}{k!}. \tag{25}$$

The predictions of $\tilde{N}$ can be improved using the information on customer orders received at each stage. The expected value ($E_t[\tilde{N}]$ for $0 \le t \le T-\tau$) can be calculated using Eq. (25) as follows:

$$E_t[\tilde{N}] = \lambda\tau. \tag{26}$$

For the case when $T \ge t > T-\tau$, $E_t[\tilde{N}]$ can be calculated as follows:

$$E_t[\tilde{N}] = q + \lambda(T-t), \tag{27}$$

where q is the number of orders received in stage $T-\tau+1$ to t.

Customer orders are fulfilled by successful production of the specified end-products, which can only be done if the required amounts of various raw materials are in stock. The raw material required can be computed almost instantly (2 days) after customer orders are received. However customer orders often vary in size and due dates, and although the order size in the near future can be closely predicted by managers with good relationships with the customers, accurate long-term predictions are still difficult to make. So for certain raw materials, rebar being a good example, the amount needed to serve customer orders at a future point in time is variable, but the variability will reduce as that point in time approaches. If we let $\tilde{d}$ be the expected (average) amount of rebar needed to serve the orders of a single customer in stages $T-\tau+1$ to T, the reduction in variability can then be expressed as:

$$E_t[\tilde{d}] \sim N(d_t, \Sigma_t) \tag{28}$$

and

$$\Sigma_{t_1} \ge \Sigma_{t_2}, \text{ when } t_1 \le t_2 \tag{29}$$

Since the material demand can be computed almost instantly upon receiving customer orders, the planned rebar supply quantities, $\tilde{D}$, required in stages $T-\tau+1$ to T can be estimated as follows:

$$\tilde{D} = \tilde{N} \cdot \tilde{d}. \tag{30}$$

The value of $E_t[\tilde{D}]$ at stage t can then be calculated as:

$$E_t[\tilde{D}] = \begin{cases} d_t \lambda \tau, & 0 \leq t \leq T - \tau \\ d_t \left(q + \lambda(T-t)\right), & T - \tau < t \leq T. \end{cases} \tag{31}$$

4 Case Study Example

After analyzing the historical price data (weekly observations of prices for rebar futures contracts in Shanghai Futures Exchange from 03/2009 to 03/2011), the parameters of the Kalman filter [13] are estimated as shown in Table 1.

Table 1. Estimation of the price model parameters

Parameter	Description	Estimate	Standard Error
κ	Short-term mean-reversion rate	1.0837	0.0132
σ_χ	Short-term volatility	0.3241	0.0056
λ_χ	Short-term risk premium	0.121	0.0126
μ'	Equilibrium drift rate	0.0095	0.0005
μ	Equilibrium risk-neutral drift rate	0.0213	0.0024
σ_ω	Equilibrium volatility	0.2976	0.0112
ρ	Correlation in increment	0.6132	0.0466

There are 6 months in the planning horizon, corresponding to 6 stages in the model. The initial price of rebar is set at 4,500 RMB per tonne and the budgeted unit price is set at 5,000 RMB per tonne, which corresponds to a total procurement budget of $5000 \times E(\tilde{D})$ Yuan. According to the purchasing manager of the company, a customer order arrives every two weeks and raw material is procured every month, so $\lambda = 4.3$ per month and $\tau = 1$. The expected demand is assumed to take unit values, i.e., $d_0 = 1$. This assumption is acceptable because the purpose of the Monte Carlo simulation experiments is to analyse risk and volatility in relative terms, and no generality is lost by such an assumption as long as the volatility Σ_t applies equally to the two strategies that are being compared, i.e. hedged and unhedged. From the historical data for rebar quantities required by the company, it is estimated that $\Sigma_0 = 0.53^2$. It then follows that:

$$\Sigma_t = \frac{T-t}{T} \Sigma_0, \quad \forall\, 0 \leq t \leq T-1. \tag{32}$$

To obtain the heuristic solution, the parameter also needs to be estimated using Eq. (13); resulting in $\alpha = 1.21 \times 10^{-3}$.

Finally, with Eq. (18) together with the price and stochastic demand updating processes, the heuristic multistage hedging model is completely defined for the specific rebar case example. Monte Carlo simulation is used to generate a large sample (5,000) of the rebar price and of the demand. The performance of the proposed multistage hedging strategy is then evaluated in terms of the expected budget overrun (EBO, see Eq. (6) and the standard deviation (SD). The results are summarised in Table 2.

Table 2. Performance evaluation of different strategies (RMB)

	SD	EBO
No hedge	1,482.70	1,038.20
Multistage strategy	1,176.30	701.5

Table 2 shows that both the SD and EBO are significantly reduced by the multistage financial hedge. The significant reduction of SD (from 1,482.7 to 1,176.3) indicates that the stability of the overall procurement spending is improved by the multistage hedge. Further, the substantial reduction in EBO (from 1,038.2 to 701.5) suggests that the procurement budget is much less violated under the multistage hedge than that with no hedging. Therefore, it can be safely concluded that the multistage hedging strategy is effective in procurement risk mitigation from the procurement budget perspective. Moreover, the percentage reduction in EBO for the multistage hedge is larger than SD; EBO is reduced by 32.4% while SD is reduced by 20.6%, see Table 2. These reductions suggest that the proposed multistage hedging strategy is more effective in reducing overspending of the procurement budget than in reducing the uncertainty of the overall total procurement spending.

A sensitivity analysis is conducted to judge the effectiveness of the multistage hedging strategy when the price and demand become highly volatile, and also when their volatility can be regarded as low. The sensitivity analysis first examines the influence of changing volatility in demand. This is done by assessing the performance of the hedge with different levels of demand volatility, i.e., by increasing/decreasing the standard deviation Σ_0 through a *changing ratio*, such that:

$$SD_{\text{changed}} = SD_{\text{initial}} \times (1 + \text{changing ratio}). \tag{33}$$

The influence of changing price volatility, represented by both the short-term volatility σ_χ and the long-term equilibrium volatility σ_χ, is also investigated. Here these two volatilities are simultaneously increased or decreased by a ratio percentage. With either demand or price volatility changing by a ratio ranging from −70% to 70%, the effects of such changes on the hedging performance are shown in Fig. 2 (demand volatility) and Fig. 3 (price volatility).

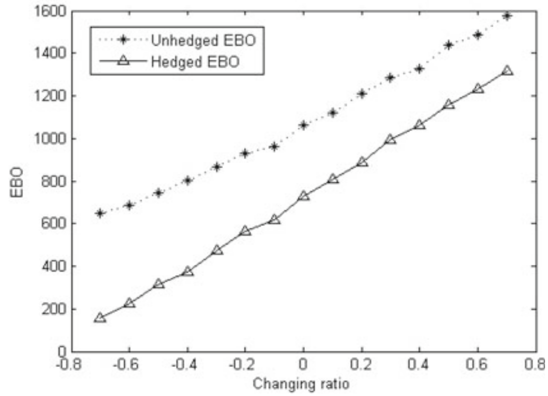


Fig. 2. Sensitivity analysis of demand volatility

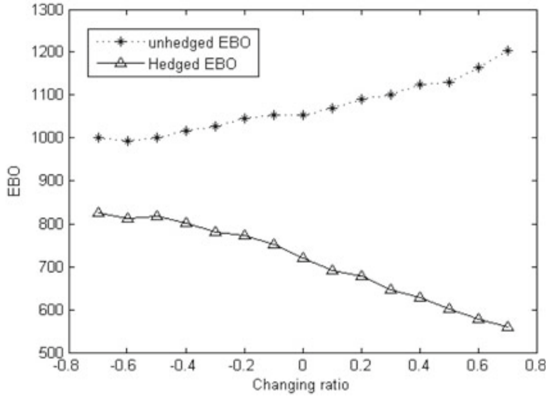


Fig. 3. Sensitivity analysis of price volatility

Figure 2 shows that changing the volatility of demand will affect the performance of the multistage strategy, an obvious result. However, the hedged EBO remains below the unhedged EBO for all changes in Σ_0, suggesting that the effectiveness of the multistage hedging strategy is always superior. When Σ_0 decreases, the hedged EBO is reduced by a larger ratio than the unhedged EBO, indicating that the procurement budget is less likely to be violated. This suggests that the multistage hedging strategy becomes relatively more powerful as Σ_0 decreases. On the other hand, the reduction ratio of the hedged EBO when compared to the unhedged EBO will decrease as Σ_0 increases. So, the multistage hedging strategy will be relatively less powerful as increases, an obvious result, since any procurement policy can be expected to deteriorate as volatility increases.

Figure 3 shows that the performance of the multistage strategy is affected by the price volatility, also an obvious result. On one hand, the multistage hedge will become relatively more powerful as σ_χ and σ_ω increase, as shown by the widening difference between the hedge and the unhedged EBO. While on the other, the multistage hedge becomes less powerful when σ_χ and σ_ω decreases, as shown by the narrowing difference between the hedged and unhedged EBO. Furthermore, Fig. 3 shows that the difference between the hedged and unhedged EBO will reduce sharply when the price volatility decreases significantly. This indicates that price volatility plays a major role in the performance of the multistage hedge model.

5 Concluding Comments

The financial hedging problem presented in this paper is modelled as a multistage strategy, and data from a machinery company is used to analyse the expected budget overrun risk in procuring rebar. A closed-form heuristic solution is obtained for the model when the raw material price is independent of customer demand. The novelty of the proposed multistage hedging strategy is that it addresses the importance of budget control, and hence the risk of possible overspending in commodity procurement is mitigated. Another distinctive feature of the model is the rebalancing of commodity futures position at the intermediate stages of the planning horizon. This rebalancing allows appropriate adjustments of the commodity futures position to be made at various stages, by taking into account new trends in the stochastic price movement, and up-to-date changes in the prediction of amount required.

The effectiveness of the model is tested via a case study in the supply of rebar for a machinery company. The results confirmed the usefulness of the proposed multistage hedging strategy in reducing procurement risk arising from variable finished product demand and raw material prices when using a predetermined level of budgetary control. This model developed in this study can be extended in several directions. First, it could be beneficial to investigate the use of other types of financial derivatives, such as options contracts, as the instrument of the hedge. Second, the multistage hedging approach could also be expanded to mitigate other types of operational risks, such as inventory risk and currency risk. Another possible extension is to use a penalty (cost) function for budgetary control, as under-spending by a certain amount is generally less serious than overspending by the same amount.

Acknowledgement. This research was supported by the National Natural Science Foundation of China (Grant No. 71601159; 71203185 and 71620107005) and the project of Research Center for System Sciences and Enterprise Development (Grant No. Xq16C03).

References

1. Alizadeh AH, Nomikos NK, Pouliasis PK (2008) A markov regime switching approach for hedging energy commodities. J Bank Financ 32(9):1970–1983
2. Caldentey R, Haugh M (2006) Optimal control and hedging of operations in the presence of financial markets. Math Oper Res 31(2):285–304
3. Chod J, Rudi N, Van Mieghem JA (2010) Operational flexibility and financial hedging: complements or substitutes? Manage Sci 56(6):1030–1045
4. Choi TM, Li D, Yan H (2003) Optimal two-stage ordering policy with bayesian information updating. J Oper Res Soc 54(8):846–859
5. Doege J, Schiltknecht P, Lüthi HJ (2006) Risk management of power portfolios and valuation of flexibility. OR Spectr 28(2):267–287
6. Eppen GD, Iyer AV (1997) Backup agreements in fashion buying-the value of upstream flexibility. Manage Sci 43(11):1469–1484
7. Lence SH (1995) On the optimal hedge under unbiased futures prices. Econ Lett 47(3–4):385–388
8. Mathews KH, Holthausen DM (1991) A simple multiperiod minimum risk hedge model. Am J Agric Econ 73(4):1020–1026
9. Mcmillan DG (2005) Time-varying hedge ratios for non-ferrous metals prices. Res Policy 30(3):186–193
10. Myers RJ, Thompson SR (1989) Generalized optimal hedge ratio estimation. Am J Agric Econ 71(5):858–868
11. Ni J, Chu LK et al (2012) A multi-stage financial hedging approach for the procurement of manufacturing materials. Eur J Oper Res 221(2):424–431
12. Ni J, Chu LK, Yen BPC (2015) Coordinating operational policy with financial hedging for risk-averse firms. Omega 59:279–289
13. Øksendal B (2006) Stochastic differential equations - an introduction with applications. Universitext 51(10):1731–1732
14. Sayın F, Karaesmen F, Özekici S (2014) Newsvendor model with random supply and financial hedging: utility-based approach. Int J Prod Econ 154(4):178–189
15. Schwartz ES (1997) The stochastic behavior of commodity prices: implications for valuation and hedging. J Financ 52(3):923–973
16. Serel D (2007) Capacity reservation under supply uncertainty. Comput Oper Res 34(4):1192–1220
17. Serel DA, Dada M, Moskowitz H (2001) Sourcing decisions with capacity reservation contracts. Eur J Oper Res 131(3):635–648
18. Shi Y, Wu F et al (2011) A portfolio approach to managing procurement risk using multi-stage stochastic programming. J Oper Res Soc 62(11):1958–1970
19. Spinler S, Huchzermeier A, Kleindorfer P (2003) Risk hedging via options contracts for physical delivery. OR Spectr 25(3):379–395
20. Xu Y (2006) Procurement risk management using commodity futures: a multistage stochastic programming approach/. Int J Prod Res
21. Zhao L, Huchzermeier A (2015) Operations–finance interface models: a literature review and framework. Eur J Oper Res 244(3):905–917

Ubiquitous Healthcare and Ubiquitousness of Chronic Disease Prevention and Control: Theory and Design

Zhihan Liu(✉)

School of Public Administration, Central South University,
Changsha 410083, People's Republic of China
liuzhihan@csu.edu.cn

Abstract. With the gradual maturing of technologies such as cloud computing, high-speed mobile Internet and Internet of things, to realize the Ubiquitous Healthcare in the health field is not only an irresistible trend, but also a realistic possibility. According to the chronic disease management theory and the service-oriented architecture, this study proposed the designing scheme and implementation methods of a mobile service platform for the big data collection and analysis of chronic disease based on cloud computing. With this platform, community residents can enjoy the convenient and quick health self-management services, while community healthcare staff can provide timely and efficient services to the patients with chronic diseases or those potentially high-risk individuals precisely. This paper is aimed to provide theoretical reference for the ubiquitous healthcare research specifically in chronic disease prevention and control, and also help in the practice of upgrading community health service and promoting a supply-side structural reform in chronic disease management.

Keywords: Big data · Chronic disease · Cloud computing · Health management · Ubiquitous healthcare

1 Introduction

Community healthcare service was identified by the World Health Organization (WHO) as an effective measure for chronic disease prevention and control. However, since China started the reform of urban healthcare system and the development of community healthcare service in 1997, the condition of suffering from chronic diseases has not been improved as much as expected and the medical expenses have also risen rapidly. Why is that the community healthcare service in China did not have satisfactory effects on residents' health? In addition to the service object factors (e.g., people's health consciousness, lifestyle and stereotype of treatment selection) and the policy factors (e.g., government investment and allocation of health resources), the traditional form of community healthcare service, the content homogenization with hospitalization service,

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_125

and the maladaptation to the needs of patients with chronic diseases are also worthy of consideration.

In fact, it is difficult to achieve the goal of building an integrated system of prevention and control towards the growing outbreak of chronic diseases just through the traditional medical information management. In 2002, Kirn [6] formally put forward the concept of Ubiquitous Healthcare (u-Healthcare or u-Health). It is a kind of information system configuration enabling individual consumers to access any types of health services through mobile computing devices whenever and wherever [1–3,5,8–13]. This concept actually originated from the idea of Ubiquitous Computing (ubicomp or Pervasive Computing) in the field of information science and technology. Transcending the traditional desktop computing, ubiquitous computing is a bran-new computing mode aimed at fusion of cyberspace and physical space, under which people could acquire the digital services freely and transparently [7].

For being humanistic, ubiquitous computing was doomed to accomplish great deeds in the fields of medicine and healthcare (especially in the health management area) which are closely related to human health and well-being, since the day of its birth. As the 10th International Conference on Ubiquitous Healthcare held in Yokohama, Japan in 2013 declared: “Enhancement in the welfare for future requires the change to the current healthcare system. Our concern for the healthcare is shifting from ‘recovery from illness’ to ‘maintaining wellness and improving quality of life’. For the care of daily health level we need special kinds of methods and technologies that we can be applied into our daily life smoothly” [10]. Many developed countries have established relatively complete u-Health systems with the following typical architecture [11] (see Fig. 1). Based on the development of Internet of Things (IoT) and cloud computing, the medical industry in Taiwan, China, has also witnessed the transition from “e-Health” to “m-Health” and eventually to “u-Health”, in a short span of ten years. However, due to the late start of ubiquitous computing research in mainland China, previous work on u-Health (especially for the chronic diseases management) is quite limited at this time [12,13]. To addresses the problem, this study is the first one in mainland China to explore the application of u-Health in the chronic disease scenario. Not only the theory and the functions were introduced, but also the technical route and the implementation methods were presented, aiming at helping the Chinese government implement its comprehensive strategy for chronic disease prevention and control.

2 Functional Requirements of Mobile Community Health Management Service Platform for Chronic Diseases Based on Cloud Computing

Chronic Disease Management (CDM) includes health information collection, health data storage, health risk assessment, chronic disease monitoring, health interventions, etc. All these tasks provides natural basis for the functional unit division of a mobile community health management service platform for chronic

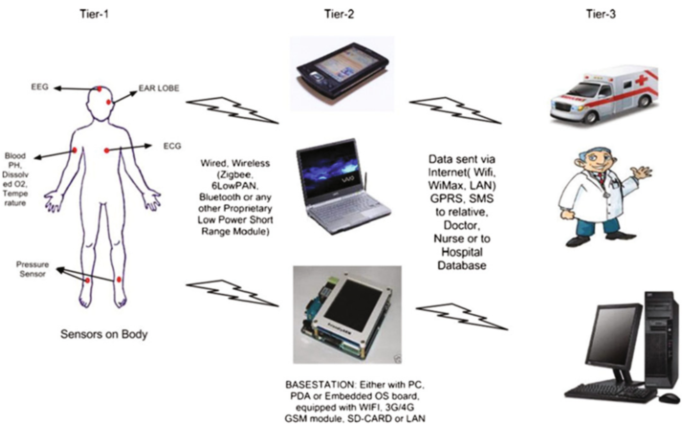


Fig. 1. Typical architecture of u-Health system

diseases. It is important to note that this study adopted the theory of Yang [14] which insists that the receptors of community prevention and control should not only include the patients with chronic diseases, but also the high-risk groups and even the healthy ones, from the perspective of community health management. Thus, the application system of this platform can be divided into two subsystems: community health service center and community residents (not just the patients). The main goal of the system is moving forward the "strategic pass" of CDM to help in cost control under the guidance of health data real-time acquisition, and carrying out a systematic, coordinated and integrated strategy on chronic disease prevention and control by in-depth health data exploring. The specific steps include: (a) collecting, managing and analyzing the personal health information used in health risk factors assessment, tracking, and health behavior guidance, for dynamical monitoring of community residents' state of health; (b) applying health interventions to the community residents based on the evaluation results; and (c) tracking and evaluating the effects of interventions (see Fig. 2).

3 Implementation Methods of Mobile Community Health Management Service Platform for Chronic Diseases Based on Cloud Computing

3.1 Technical Route

The system adopted the MySQL database for data storage, and used Java language for realizing the background management function and the apps development based on Android and IOS. Given the complexity of apps on the community health management service platform, the platform was designed in according to Service-Oriented Architecture (SOA), providing functions such as service definition, development, deployment, and running, etc. Logic reuse was implemented

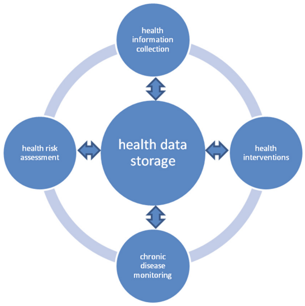


Fig. 2. System functional requirements of the service platform

through the loose coupling between the layers, while the loose coupling between modules on the same layer was implemented through the form of services. All of these equipped the platform with good ability of extension.

3.2 Overall Architecture

The technical implementation of the service platform function involved various kinds of data acquisition sensors, data transferring networks, storage databases, cloud computing services and apps for data processing, interaction with users, and many other complex factors. In order to improve the robustness of the program and make the design more reasonable, this study adopted a five-layer architecture (see Fig. 3).

Each layer has a clear job assignment as follows:

(1) Perception layer (PL): To be in charge of the perception, identification, collection and processing of residents' health data (under automatic control), and connecting data acquisition devices with network layer and application layer through communication module;
(2) Network layer (NL): To realize the function of information transmission (i.e., uploading collected data and downloading information from application layer), routing, and control. This layer is made up with core network and access network, and depend on the public telecommunication networks and the Internet;
(3) Support services layer (SSL): To set up a basic support service platform for the processing of CDM related information and operations. Specifically, it at least covers a SOA platform, an integration service platform for mass data, a cloud computing service platform, a communication service platform, and an IoT service platform;
(4) Application layer (AL): To carry and realize the various function of community health management, i.e., health risk assessing, chronic disease monitoring, health intervention and outcome evaluation;
(5) Interaction layer (IL): To provide users with a variety of interactive terminals (such as PC, mobile phones, etc.), realize the interaction of information and instructions between people and system.

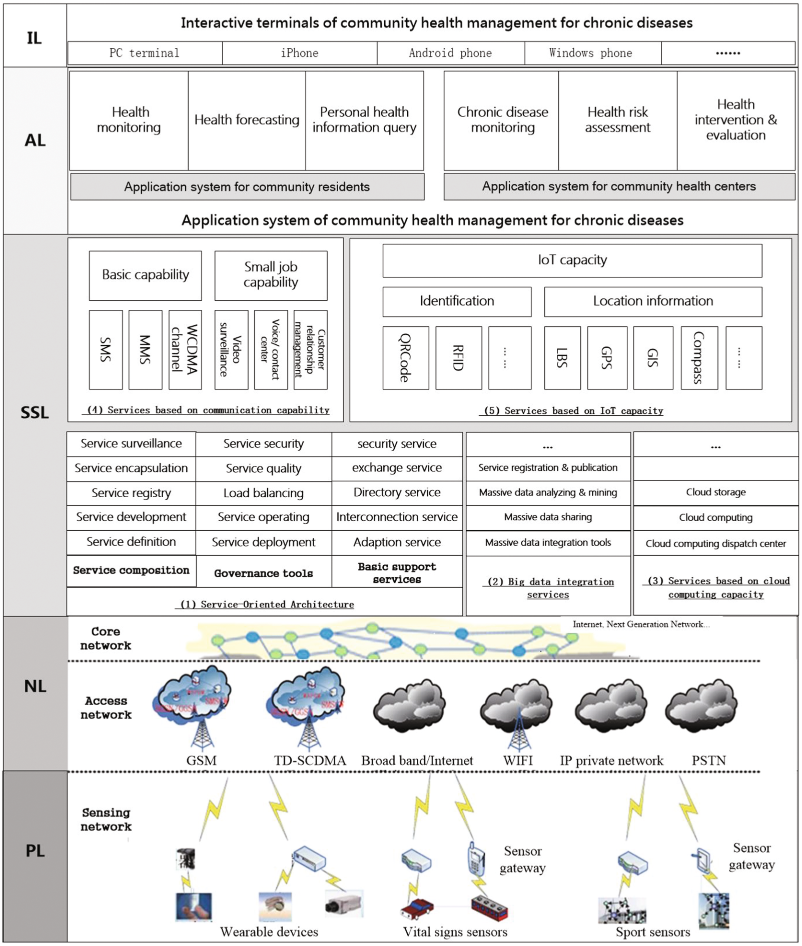


Fig. 3. Overall architecture of the mobile service platform for community health management of chronic disease

3.3 Public Service Modules

Four public service modules were designed uniformly as the system implementation process may require cloud computing, massive data, IoT, communication and other public services on many aspects.

(1) Big data integration services: According to the characteristics of massive structured data and unstructured data in the vast distributed network environment, the platform provided big data integration services based on SOA, including organization, storage, sharing, release, analysis and mining, etc.
(2) Cloud computing services: Considering ubiquitous healthcare are higher than current Internet several orders of magnitude on the amounts of data storage and operation, a cloud computing solution was needed to be integrated

into the platform design on the basis of traditional IT system, to improve the computing ability. For the community healthcare information centers, cloud computing could make full use of the great deal of idle computing and storage capacity, and provide important technical means for adaption to the flexibility of business growth and reduction of application deployment cost.

(3) Services based on IoT capacity: By accessing to a variety of IoT terminals and applications through standardized protocols and various adapter, the platform was provided with data acquisition and transmission capacity of IoT terminals, and finally realized the interaction and cooperation between machine and machine, and between machine and people.

(4) Services based on communication capacity: This module included the communication capacity required, such as SNS, WeChat, positioning, and call centers, etc. The platform may also need some other third-party services and resources.

4 Conclusion

Just as Dr. Robert M. Kaplan, who was Associate Director for Behavioral and Social Sciences in the National Institutes of Health of US at that time said in his paper: “Health-related information collected in psychological laboratories may not be representative of people’s everyday health. For at least 70 years, there has been a call for methods that sample experiences from everyday environments and circumstances” [4]. Ubiquitous healthcare not only provides us with this method, but also with the development direction of medical information construction. Under the u-Health theory and aiming at the key requirements and typical applications of CDM, this study is the first in mainland China to present a mobile service platform for the big data collection and analysis of chronic disease, on the comprehensive basis of cloud computing, intelligent sensing and perception technologies of IoT, data acquisition, storage and processing technologies, and computer decision support technology. The platform can be used in the health data acquisition, supervision and forecast for community residents, contributing to improve the community health services and lift the population health level. It is expected to help in realizing the humanistic self-health management of community residents, and providing community health centers with the ubiquitous, comprehensive prevention and control services of chronic diseases, so as to accelerate the supply-side structural health reform in the field of CDM. In consideration of the rapid progress in the international development of information technology, direction for future research would be getting the platform put into practice as soon as possible, and adjusting and improving the system in a timely manner according to the process of Chinese healthcare system reform.

Acknowledgement. This study was supported by National Natural Science Foundation of China (grant No. 71603289), China Postdoctoral Science Foundation (grant No. 2016M590759), Social Science Fund of Hunan Province (grant No. 15JD61), and Postdoctoral Science Foundation of Central South University (No. 153656). The author declares no conflicts of interest.

References

1. Arnrich B, Mayora O, Bardram J (2010) Pervasive or ubiquitous healthcare? Methods Inf Med 49:65–66
2. Conejar RJ, Kim HK (2015) Designing u-healthcare web services system. Int J Softw Eng Appl 9(3):209–216
3. He C, Fan X, Li Y (2013) Toward ubiquitous healthcare services with a novel efficient cloud platform. IEEE Trans Biomed Eng 60(1):230–234
4. Kaplan RM, Stone AA (2013) Bringing the laboratory and clinic to the community: mobile technologies for health promotion and disease prevention. Ann Rev Psychol 64:471–498
5. Kim J, Ahn CW (2011) Diabetes management system based on ubiquitous healthcare. J Paediatr Child Health 12(3):133
6. Kirn S (2002) Ubiquitous healthcare: The OnkoNet mobile agents architecture. In: Objects, Components, Architectures, Services, and Applications for a Networked World. Springer, Heidelberg, pp 265–277
7. Krumm J (2010) Ubiquitous computing fundamentals. Ergonomics 53(5):724–725
8. Panagiotakopoulos T, Fengou MA et al (2009) Ubiquitous healthcare. In: Biocomputation and Biomedical Informatics: Case Studies and Applications, p 254
9. Shin D, Shin D, Shin D (2016) Ubiquitous healthcare platform for chronic patients. In: International Conference on Platform Technology and Service, pp 1–6
10. Tamura T, Park KS (2016) Invitation to the 10th International Conference on Ubiquitous Healthcare (u-healthcare 2013). http://u-healthcare2013.l-bmi.org/
11. Touati F, Tabish R (2013) U-Healthcare system: state-of-the-art review and challenges. J Med Syst 37(3):1–20
12. Wu X, Ye M et al (2012) Pervasive medical information management and services: key techniques and challenges. Chin J Comput 35(5):827–845
13. Xie Y (2016) Ethical perspectives of ubiquitous healthcare. J Jishou Univ (Social Science Edition) 37(3):56–62 (in Chinese)
14. Yang JX (2010) Thoughts about strategies for chronic disease integrated control on the basis of community health management. Chin Health Econ 29(7):67–69 (in Chinese)

Can Media Supervision Improve the Quality of the Firms Internal Control?

Tian Yang[1,2], Qianwei Ying[3(✉)], and Xingyao Li[3]

[1] The College of Literature and Journalism, Sichuan University, Chengdu 610065, People's Republic of China
[2] School of Overseas Education, Sichuan University, Chengdu 610065, People's Republic of China
[3] Business School, Sichuan University, Chengdu, People's Republic of China
yingqw@scu.edu.cn

Abstract. Using the sample of firms listed in China Growth Enterprise Market (GEM) from 2011 to 2015 we empirically test the effect of media coverage on the corporate internal control quality and the quality of internal control information disclosure. We found that the higher the degree of media coverage would lead to higher quality of internal control and information disclosure. The empirical results of this paper indicate that media can act as an important external supervision mechanism besides of state laws and regulations to regulate the behavior of corporate internal control, to protect the interests of investors, and to enhance investors' confidence.

Keywords: Media coverage · Internal control · Information disclosure

1 Introduction

With the development of information technology and the spread of internet, media's functions as supervision and governance are taking more and more significant roles in economy. Media, regarded as a forth power being independent of legislation, administration and judiciary, plays an essential role in disseminating information, restraining incompliant actions of government and businesses, and enhancing market's efficiency. Increasing number of scholars in economics and management have focused on studying the corporate governance role of media as well as its influence mechanism. Dyck and Zingales [8] found that the pressure from media supervision contributes drastically to reducing the personal interests of the controlling shareholders from their controlling rights. Dyck and Zingales [9] discovered that media supervision is able to effectively suppress or modify the corporate decisions that may harm external investors' interests. Miller [20] made a conclusion that through rebroadcasting investigation of information from information intermediary like securities analysts or original ones, media can disclose accounting frauds in advance. Joe et al. [15] claimed that an inefficient board would take actively measures to improve efficiency after exposed by

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_126

media. Maistriau and Bonardi [19] reported that media visibility has a positive effect on CSR performance of a sample of British firms, and Zyglidopoulos et al. [22] found that media coverage is positively associated with heightened levels of CSR performance of a sample of S&P 500 firms.

Using a large cross-country sample of firms, Ghoul et al. [10] further found strong evidence that firms engage in more CSR activities if located in countries where the media has more freedom. Dai et al. [5] investigated whether the media plays a role in corporate governance by disseminating news. Using a comprehensive data set of corporate and insider news coverage for the 2001–2012 period, they showed that the media reduces insiders' future trading profits by disseminating news on prior insiders' trades available from regulatory filings. Based on a sample of over two million newspaper articles, Kim et al. [17] found the media in China has an incremental impact on stock price efficiency, and a market-driven media can play the role of compensating for the underdeveloped governance institutions in transitional economies such as China.

Although prior studies accumulate evidence that supports the monitoring role of the media, they ignore an important issue: Can media supervision encourage firms to take precautions and initiate to improve the quality of internal controls and information disclosures? Previous literature regards media supervision as a substitute of internal controls in corporate governance, focuses on corporate inherent characteristics only when investigating determinants of corporate internal controls, and neglects the potential interaction between media's external supervision and corporate internal controls. Internal controls mean a series of restrictive organizing, planning, procedures and methods executed in entities, aiming to boost operating efficiency, acquire and use resources in an effective manner, and achieve given objectives. The setting-up, perfection and valid operation of business internal controls have direct impact on disclosures of financial information and investors' interests. Therefore, launching a research on the way media supervision influences corporate governance has theoretical and practical meanings. This article selects data of GEM firms in China's stock market from 2011 to 2015, and studies impacts of media coverage on internal controls and related information disclosures. The contributions of this paper mainly lie in the following two aspects.

Firstly, this paper enriches the literature on the corporate governance role of media. Existing studies on the corporate governance role of media mainly focus on the effect of media coverage on firms' socially beneficial behaviors such as CSR performance, or on firms' frauds and misconducts hurting the interests of the investors. This paper extends the scope of this research area by investigating the effect of media coverage on firms' internal control quality and internal control information disclosure.

Secondly, this paper also enriches the literature on the determinants of internal control quality and disclosure. Most prior studies on the determinants of internal control quality and disclosure mainly focus on corporate internal characteristics (e.g., [4,7,11]), while very few researches investigated the role of external factors, such as auditor expertise (e.g., [12,21]). In this paper, we find that media coverage, as another important external governance mechanism, also contributes significantly to the firms' internal control quality and disclosure.

2 Prior Literature and Hypotheses Development

As an important segment in corporate governance, the quality of internal controls plays a key role in business operation and development, as well as the protection of the interests of investors. Existing research indicated that good internal controls would suppress earning management [1], and reduce financing cost of firms effectively [16].

Most previous literature about the determinants of internal control quality mainly focused on corporate internal characteristics. Doyle et al. [7] examined determinants of weaknesses in internal control for 779 firms disclosing material weaknesses from August 2002 to 2005. They found that these firms tend to be smaller, younger, financially weaker, more complex, growing rapidly, or undergoing restructuring. Firms with more serious entity-wide control problems are smaller, younger and weaker financially, while firms with less severe, account-specific problems are healthy financially but have complex, diversified, and rapidly changing operations. Guo et al. [11] investigated the role of employment policies in reducing internal control ineffectiveness and financial restatements. They found that employee-friendly policies significantly reduce the propensity for employee-related material weaknesses. Moreover, they illustrated that financial restatements caused by unintentional errors are less likely to arise in firms that invest more in employee benefits. Chen et al. [4] documented a negative relation between board independence and the disclosure of internal control weaknesses. They also documented that the negative relation is stronger for firms with unitary leadership (combined positions of CEO and Chairman) than for firms with dual leadership.

Only a few studies studied the effects external factors on internal control, and these studies mainly focus on the role of audit expertise. Haislip et al. [12] argued that greater audit firm IT expertise promotes improved internal controls for their clients, especially those controls that are dependent on IT. Schroeder and Shepardson [21] found that internal control audits initially provided internal control quality benefits. However, after the 2007 auditing standards change, internal control quality deteriorated for ICFR audited versus unaudited firms.

An important external governance mechanism, media supervision, has been ignored by prior studies on the determinant of internal control quality. Most countries, particularly the transitional countries, are facing problems of imperfect legal system and inadequate protection of the investors' interests. Among the alternative governance mechanisms besides of legal system, media supervision is one of the most important. Most scholars hold an optimistic opinion towards the role of media supervision playing in corporate governance. Dyck and Zingales [8] found that media supervision could reduce the personal interests of the controlling shareholders on the sacrifice of other shareholders. Miller [20] showed that the media plays a positive role in the process of revealing accounting scandals.

Dyck and Zingales [9] further found that the media exposure could encourage enterprises to correct violations of the rights and interests of external investors. Zyglidopoulos et al. [22] found that media coverage is positively associated with heightened levels of CSR performance. Dai et al. [5] showed that media coverage can reduce insiders' future trading profits by disseminating news on prior insiders' trades available from regulatory filings.

We can conclude that existing researches have provided sufficient evidences for the corporate governance role of media supervising. Therefore, we ascertain that media coverage can urge firms to improve internal control quality, and reduce adverse actions to investors as far as possible, to avoid greater negative market response after exposed by media. At the same time, media coverage also encourages firms to build a good image in external stakeholders and enhance business values through improving internal controls. Hence, we can obtain the following hypothesis:

Hypothesis 1: Media coverage has a positive effect on the firms' quality of internal controls.

One important characteristic of internal controls is that they consist of internal managerial processes that are not directly observable for investors. Investors' judgment of the effectiveness of these internal controls is based on the information disclosed by managers.

From shareholders' perspective, higher levels of internal control disclosure reduce investors' information risk, lower down their required rate of return (e.g., [3,6]). At the same time, higher levels of disclosure enable shareholders to monitor managers more closely, resulting in fewer agency problems.

From the manager's perspective, the decision to voluntarily disclose information is based on a tradeoff between the expected benefits and costs of disclosing [14]. The main benefit of voluntarily disclosing information on the firm's internal controls is that it adds to managerial reputation building. Not having a reputation for credible reporting not only reduces the effectiveness of the manager's communication efforts, but also adversely affects her reputation in the managerial labor market [18]. Establishing a reputation for credible reporting requires disclosure of accurate and timely information as well as of information that is complete [14,20]. This means that career concerns incentivize a manager to voluntarily disclose information, even if the information is not favorable [3]. Hammersley et al. [13] find a significantly negative abnormal return following a firm's announcement that the firm's internal controls were not effective, but more importantly also document that the adverse effects on returns are more pronounced when the firm's managers claim that the internal controls are effective but the independent auditor report indicates that they were not.

An important cost of voluntarily disclosing information on the firm's internal controls is that it may have potential legal consequences. If a manager discloses inaccurate information or if a manager discloses incomplete information managers could be sued, face legal liability, and owe damages [3]. Moreover, once this matter is known to the public, it will also adversely influence the manager's reputation on the managerial labor market. To guarantee the accuracy and quality of the information disclosures on internal controls, the manager might need to pay extra efforts and related costs.

The above analysis and discussion on the benefits and costs of disclosing internal controls information implicate the importation role of the managers' reputation concerns. We can expect that the reputation mechanism will even be more important when a firm is in the spotlight of media coverage, since media can spread information to magnify the influence of signals and scope of audience, drawing more investors' attention [2]. When the media coverage is higher, the reputational benefit of disclosing high quality information of internal control will be larger, and the reputational cost of concealing information or disclosing poor information on internal controls will be greater. Therefore, higher media coverage can lead to higher quality of internal control information disclosure. We can conclude the second hypothesis as follows:

Hypothesis 2: Media coverage has a positive effect on the quality of internal control information disclosure.

3 Research Design

3.1 Data and Variables

This article selects GEM Growth Enterprise Market firms from 2011 to 2015 as initial sample, and eliminates observations with missing data on the key variables of this paper. The reason why we choose GEM firms is that these firms, which have small sizes and young ages in most cases, are opaque in information disclosure and have higher business risks in operation. In this case, the quality of internal controls and information disclosure of GEM firms are of great importance for investors to understand their business situations. Besides, this article uses balanced panel data, keeping only the firms with public financial data in each year from 2011 to 2015. The final sample of this paper includes 151 GEM firms, 755 observations in the period ranging from 2011 to 2015.

The corporate internal control quality index and its information disclosure quality index derive from the DIB Internal Control and Risk Management Database. Higher values of the two indexes represent higher qualities of the internal control and information disclosure respectively. The data of media coverage is from the News and Reports Database of WIND Info, which records the Chinese listed firms' daily news reported by more than 100 major financial newspapers and websites in China, covering nearly all the relevant news about public firms from major financial media in China. Using this database, we count the annual number of news reports for each GEM listed firm, and construct a media coverage indicator by taking the natural logarithm of one plus the number of new reports.

Referring to existing research results about determinants of internal control quality and information disclosure quality (e.g., [7]), our paper includes a set of basic financial characteristics and corporate governance variables as control variables. The basic financial characteristics variables include the firm size (natural logarithm of total assets), asset-liability ratio, return on assets, and sales growth.

The corporate governance variables include ownership concentration (the total share ratio of the top 10 major shareholders), executives' ratio of shareholding, equity separation degree (the quadratic sum of the share ratio of the top five major shareholders), and the separation of ownership and control. In addition, we also control the industry dummies. The data of these control variables is from CSMAR database. Continuous variables are censored at the 1st and 99th percentiles to mitigate the effect of extreme values. The descriptive statistics of variables are shown in Table 1.

Table 1. Descriptive statistics of variables

Na	Description	Mean	S.D.	Median	Min	Max
IC	Internal control quality index	6.833	0.405	6.891	5.165	7.524
Disclosure	Information disclosure of Internal control index	0.382	0.098	0.37	0.18	0.58
Media	Media concern extent	2.58	0.837	2.477	1.097	4.628
Size	Corporate size	20.785	0.528	20.77	19.65	22.25
LEV	Asset-liability ratio	0.171	0.123	0.142	0.028	0.619
ROA	Return on Assets	0.07	0.037	0.065	−0.003	0.164
Growth	Sales growth	0.317	0.307	0.263	−0.21	1.512
Ten	Ownership concentration	0.678	0.089	0.691	0.448	0.857
Mshare	Executives' shareholding	0.305	0.239	0.358	0	0.708
Herf_5	Equity separation degree	0.16	0.087	0.142	0.022	0.392
Separation	Separation of control and ownership	3.662	6.873	3.224	0	31.216

3.2 Empirical Modelling

Based on the preceding theoretical analysis, this paper established the following model (1) and (2) to analyze the effects of the media coverage on the qualities of internal control and information disclosure of internal control respectively:

$$\begin{aligned} IC_{i,t} = &\beta_0 + \beta_1 \text{Media}_{it} + \beta_2 \text{LEV}_{it} + \beta_3 \text{ROA}_{it} + \beta_5 \text{Growth}_{it} + \beta_6 \text{Ten}_{it} \\ &+ \beta_7 \text{Herl\_5}_{it} + \beta_8 \text{Mshare}_{i,t} + \beta_9 \text{Separation}_{i,t} \\ &+ \text{Industry and year fixed effects} + \varepsilon_{it}, \end{aligned} \tag{1}$$

$$\begin{aligned} \text{Disclosure}_{i,t} = &\beta_0 + \beta_1 \text{Media}_{it} + \beta_2 \text{LEV}_{it} + \beta_3 \text{ROA}_{it} + \beta_5 \text{Growth}_{it} \\ &+ \beta_6 \text{Ten}_{it} + \beta_7 \text{Herl\_5}_{it} + \beta_8 \text{Mshare}_{i,t} + \beta_9 \text{Separation}_{i,t} \\ &+ \text{Industry and year fixed effects} + \varepsilon_{it}, \end{aligned} \tag{2}$$

where IC represents the internal control quality index. Disclosure represents the quality of internal control information disclosure, and Media represents the extent of media coverage. The definitions and descriptions of other control variables are as shown in Table 1. Meanwhile, Industry and year fixed effects are also controlled in the above regressions. This article uses the OLS regression model, and clusters standard errors of the coefficients at firm level for more accurate statistical inference [18]. β_0 is the constant term in regression, while β_1 to β_9 represent the regression coefficients of the explanatory and control variables. We are mainly interested in the coefficient β_1, which indicates the effect of media coverage on internal control quality in model (1) or internal control disclosure in model (2).

4 Empirical Results and Analyses

4.1 Regression Results

Following the above regression models (1) and (2), this paper conducts regression analyses respectively with internal control quality (IC) and its information disclosure quality (Disclosure) as dependent variables. The regression results are shown in Table 2.

In the column 1 of Table 2, we used media coverage as the sole explanatory variable, without controlling other firm characteristic variables. The result shows that the regression coefficient of media coverage is positive and significant at 10% significance level, with a value of 0.099. This regression result indicates

Table 2. The media coverage's effects on the quality of internal control and internal control information disclosure

	– 1	– 2	– 3	– 4	– 5	– 6
	IC	IC	IC	Disclosure	Disclosure	Disclosure
Media	0.099^{*}	0.161^{**}	0.169^{***}	0.146^{**}	0.168^{***}	0.153^{**}
	– 1.89	– 2.15	– 2.66	– 2.13	– 2.72	– 2.49
Size		0.297^{***}	0.137^{**}		0.135^{**}	0.131^{**}
		– 3.86	– 2.27		– 2.15	– 1.98
LEV		– 0.03	– 0.029		-0.196^{*}	– 0.17
		(– 0.31)	(– 0.30)		(– 1.94)	(– 1.49)
ROA		0.256^{**}	0.180^{**}		0.245^{**}	0.255^{**}
		– 2.39	– 1.99		– 2.11	– 2.36
Growth		0.069	0.081		0.195^{*}	0.181
		– 0.89	– 1.11		– 1.71	– 1.5
Ten			0.004			0.111
			– 0.04			– 0.96
Mshare			0.104			0.06
			– 1.14			– 0.65
Herf_5			– 0.086			0.039
			(– 0.91)			– 0.34
Separation			– 0.05			0.07
			(– 0.56)			– 0.89
constant	0.935^{***}	-0.755^{***}	0.549	1.116^{***}	-0.645^{***}	1.211^{***}
	– 18.4	(– 3.21)	– 1.39	– 23.32	(– 2.89)	– 2.71
Industry dummy	YES	YES	YES	YES	YES	YES
N	755	755	755	755	755	755
R_2	0.158	0.188	0.236	0.133	0.207	0.212

Note: ***, **, * represents significance at 1%, 5% and 10% level respectively.
The t-statistics reported in parentheses are based on standard errors clustered by firm.

that media coverage has a positive and significant effect on internal control quality. In column 2, we controlled the basic financial characteristics of the firms, including firm Size (Size), asset-liability ratio (LEV), return on assets (ROA), and the firms' sales growth (Growth). The positive regression coefficient of media coverage, with a value of 0.161, is significant at 5% level, indicating that the positive relationship between media coverage and internal control quality becomes more apparent after controlling the influences of these variables of firms' financial characteristics. In column 3, we further controlled the corporate governance characteristics, including ownership concentration (the total share ratio of the top 10 major shareholders), executives' share holding, equity separation degree, and the separation of ownership and control. The results show that after the control of the corporate governance, media coverage illustrates an even greater positive effect ($\beta_1 = 0.169$, $p < 0.1$) on the internal control quality. Thus, we can conclude that media coverage does help improve the quality of the firm's internal control, which is consistent with Hypothesis 1 above.

Column 4 to 6 use quality of internal control information disclosure as dependent variable. It could be found that whether using media coverage as an explanatory variable alone, or controlling the firm's basic financial characteristics and characteristics of corporate governance, there remains a positive effect of media coverage on the quality of internal control information disclosure. Thus, the media supervision not only helps to improve the quality of internal control of the company, but also help to improve the company's internal control information disclosure, thereby enhancing the company's transparency and self-restraint.

4.2 Robustness Test

In the above regressions, the measure of media coverage includes the news reports from the mainstream financial media channels in China, but excludes the Wind Info's own news reports. Although Wind Info is not a common source of information for retail investors, it is widely regarded by institutional investors and analysts as one of the important sources of information access. In order to test the reliability of the previous empirical results, this paper uses an alternative measure of media coverage, which includes Wind Info's own news reports, to do the robustness test. The regression results in Table 3 show that the main conclusions of the paper have not changed because of the replacement of media coverage measure. With the alternative measure of media coverage, denoted as *Media_A*, the regression results still show that media coverage has a positive effect on the quality of internal control and the quality of internal control information disclosure, and this conclusion remains the same regardless of whether we control the company's basic financial characteristics or corporate governance variables.

Table 3. The robustness test using alternative measure of media coverage

	− 1	− 3	− 4	− 1	− 3	− 4
	IC	IC	IC	Disclosure	Disclosure	Disclosure
Media_A	0.142*	0.186**	0.205**	0.129*	0.180**	0.193**
	− 1.81	− 2.08	− 2.39	− 1.71	− 2.1	− 2.46
Size		0.207***	0.133**		0.132**	0.135**
		− 3.01	− 2.15		− 2.03	− 2.18
LEV		0.042	− 0.012		− 0.190*	− 0.154
		− 0.43	(− 0.12)		(− 1.81)	(− 1.29)
ROA		0.232**	0.182**		0.248**	0.260**
		− 2.18	− 2.07		− 2.16	− 2.48
Growth		0.035	0.017		0.188	0.175
		− 0.35	− 0.17		− 1.62	− 1.43
Ten			0.003			0.105
			− 0.03			− 0.9
Mshare			0.117			0.072
			− 1.28			− 0.78
Herf_5			− 0.055			0.068
			(− 0.59)			− 0.58
Separation			− 0.085			0.093
			(− 0.91)			− 1
_cons	0.658***	− 0.767***	0.728*	1.726***	− 0.649***	1.375***
	− 8.86	(− 3.26)	− 1.75	− 19.84	(− 2.80)	− 2.93
N	755	755	755	755	755	755
R2	0.168	0.194	0.234	0.121	0.194	0.21

Note: ***, **, * represents significance at 1%, 5% and 10% level respectively. The t-statistics reported in parentheses are based on standard errors clustered by firm.

5 Conclusion

Based on the financial data of 151 listed companies in GEM from 2011 to 2013, this paper analyzes the impact of media coverage on the internal control quality and its information disclosure quality. The results show that media coverage positively relates to the quality of internal control and the quality of internal control information disclosure, and the conclusion is still valid after controlling corporate financial characteristics and corporate governance characteristics. This article provides new evidence for the corporate governance role of media supervision from the perspective of corporate internal control. The conclusion of this paper implicates that, under the imperfect Chinese legal system and market institutional environment, media, as an alternative form of supervision mechanism other than legal system, plays an important role in regulating the behavior of listed companies and protecting the interests of small and medium shareholders. In the current stage of China's market-oriented reform, strengthening the external supervision role of the media is of great significance to improving corporate governance of listed companies and protecting the interests of investors.

Acknowledgements. The authors would like to acknowledge the financial support from the National Science Foundation of China (NSFC71373167), the social science research fund of Sichuan University (skqx201608), and Excellent Young Scholars Research program of Sichuan University, (2013SCU04A32).

References

1. Ashbaugh-Skaife H, Collins DW et al (2008) The effect of sox internal control deficiencies and their remediation on accrual quality. Acc Rev 83(1):217–250
2. Barber BM, Odean T (2008) All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev Finan Stud 21(2):785–818
3. Campbell JL, Chen H et al (2014) The information content of mandatory risk factor disclosures in corporate filings. Rev Acc Stud 19(1):396–455
4. Chen Y, Knechel WR et al (2016) Board independence and internal control weakness: evidence from sox 404 disclosures
5. Dai L, Parwada JT, Zhang B (2015) The governance effect of the media's news dissemination role: evidence from insider trading. J Acc Res 53(2):331–366
6. Dan D, Hogan C et al (2011) Internal control disclosures, monitoring, and the cost of debt. Acc Rev 86(4):1131–1156
7. Doyle J, Ge W, Mcvay S (2007) Determinants of weaknesses in internal control over financial reporting. J Acc Econ 44(1–2):193–223
8. Dyck A, Zingales L (2004) Private benefits of control: an international comparison. J Finan 63(2):537–600
9. Dyck A, Zingales L (2008) The corporate governance role of the media: evidence from Russia. J Finan 63(3):1093–1135
10. Ghoul SE, Guedhami O et al (2016) New evidence on the role of the media in corporate social responsibility. J Bus Ethics 62:1–29
11. Guo J, Huang P et al (2015) The effect of employee treatment policies on internal control weaknesses and financial restatements. Soc Sci Electron Publishing 91(4):1167–1194
12. Haislip JZ, Peters GF, Richardson VJ (2016) The effect of auditor it expertise on internal controls. Int J Acc Inf Syst 20:1–15
13. Hammersley JS, Myers LA, Shakespeare C (2008) Market reactions to the disclosure of internal control weaknesses and to the characteristics of those weaknesses under section 302 of the sarbanes oxley act of 2002. Rev Acc Stud 13(1):141–165
14. Healy PM, Palepu KG (2001) Information asymmetry, corporate disclosure, and the capital markets: a review of the empirical disclosure literature. J Acc Econ 31(1–3):405–440
15. Joe JR, Louis H, Robinson D (2009) Managers' and investors' responses to media exposure of board ineffectiveness. J Finan Quant Anal 44(3):579–605
16. Kim JB, Song BY, Zhang L (2011) Internal control weakness and bank loan contracting: evidence from sox section 404 disclosures. Acc Rev 86(4):1157–1188

17. Kim JB, Yu Z, Zhang H (2016) Can media exposure improve stock price efficiency in China and why? China J Acc Res 9(2):83–114
18. Kothari SP, Shu S, Wysocki PD (2009) Do managers withhold bad news? J Acc Res 47(1):241–276
19. Maistriau EA, Bonardi JP (2014) How much does negative public exposure on environmental issues increase environmental performance? Academy of Management Annual Meeting Proceedings 2014(1):11328
20. Miller GS (2006) The press as a watchdog for accounting fraud. J Acc Res 44(5):1001–1033
21. Schroeder JH, Shepardson ML (2015) Do sox 404 control audits and management assessments improve overall internal control system quality? Acc Rev A Q J Am Acc Assoc 91:1513–1541
22. Zyglidopoulos SC, Georgiadis AP et al (2012) Does media attention drive corporate social responsibility? J Bus Res 65(11):1622–1627

Fuzzy Chance Constrained Twin Support Vector Machine for Uncertain Classification

Qilin Cao, Xiaodan Fu(✉), and Yaomin Guo

Business School, Sichuan University,
Chengdu 610065, People's Republic of China
lovesens@163.com

Abstract. In this paper, we propose a new fuzzy chance constrained twin support vector machine (FCC-TWSVM) via chance constrained programming formulation and fuzzy membership, which can deal data with measurement noise efficiently. This paper studies twin support vector machine classification when data points are uncertain statistically. With some properties known for the distribution, the FCC-TWSVM model is used to ensure the small probability of misclassification for the uncertain data. The FCC-TWSVM model can be transformed to second-order cone programming (SOCP) by the properties of moment information of uncertain data and the dual problem of SOCP model is introduced. In addition, we also show the performance of FCC-TWSVM model in artificial data and real data by numerical experiments.

Keywords: Fuzzy data mining · Support vector machine · Robust optimization · Chance constraints · Venture capital

1 Introduction

Nowadays, support vector machines (SVMs) are considered as one of the most effective learning methods for classification, which emerged from research of statistical learning theory [10,14]. The main idea of this classification technique is by mapping the data to the higher dimensional space with some kernel methods and then determining a hyperplane separating binary classes with maximal margin [9,19,26].

In recent years, SVM classification methods have made breakthrough progress and enjoyed great success in many fields. Binary data classification methods have made breakthrough progress in recent years. Mangasarian et al. [15] proposed generalized eigenvalue proximal support vector machine (GEPSVM). Jayadeva et al. [12] proposed a twin support vector machine (TWSVM) to solve the classification of binary data, motivated by GEPSVM. The main idea of TWSVM is generating two nonparallel planes, one of which is closest to one class and another plane is as far as possible from the other. At the same time, the ν-TWSVM [18] was proposed for handling outliers as an extension of TWSVM. Some extensions to the TWSVM can be founded in [8,13,20].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_127

For the methods mentioned above, the parameters in the training sets are implicitly assumed to be known exactly. However, in real world application, the parameters have perturbations since they are estimated from the data subject to measurement and statistic errors [11]. When the data points are uncertain, different have been proposed to formulate the traditional SVM with uncertainties. Bi et al. [5] assumed the data points are subject to an additive noise which is bounded by the norm and proposed a very direct model. However, this model cannot guarantee a generally good performance on the uncertainty set. To guarantee an optimal performance when the worst case scenario constraints are still satisfied, robust optimization is utilized. Trafalis et al. [22,23,25] proposed a robust optimization model when the perturbation of the uncertain data is bounded by norm. Robust optimization [4,21] is also used when the constraint is chance constraint which is to ensure the small probability of misclassification for the uncertain data. Based on different bounding inequalities, Ben Tal et al. [1,2] employed moment information of uncertain training points to developing different chance-constrained SVM model. However, to our best knowledge, there is no researcher considering the chance-constrained optimization in TWSVM problem. Therefore, it is interesting and important to study the TWSVM with chance constraints for the uncertain data classification problem. The main purpose of this paper is to make an attempt in this direction.

Combining the capability of processing the uncertainty of chance constraints and the benefits of TWSVM, in this paper, we propose a fuzzy chance constrained twin support vector machine (FCC-TWSVM). The main method of this paper is to apply the moment information of uncertain data to transform into second order cone programming (SOCP). The rest of this paper is organized as follows. Section 2 recalls SVM and TWSVM briefly. In Sect. 3, we introduce the model of FCC-TWSVM. Experimental results on the uncertain data sets are presented in Sect. 4. Conclusions are provided in Sect. 5.

2 Preliminaries

In this section, we briefly recall some concepts of TWSVM and CC-TWSVM for binary classification problem.

2.1 TWSVM

Consider a binary classification problem of l_1 positive points and l_2 negative points ($l_1 + l_2 = l$). Suppose that data points belong to positive class are denoted by $A \in \mathbf{R}^{l_1 \times n}$, where each row $A_i \in \mathbf{R}^n$ $(i = 1, \cdots, l)$ represents a data point with label +1. Similarly, $B \in \mathbf{R}^{l_2 \times n}$ represents all the data points with the label −1. The TWSVM determines two nonparallel hyperplanes:

$$f_+(x) = w_+^T x + b_+ = 0 \quad and \quad f_-(x) = w_-^T x + b_- = 0, \tag{1}$$

where $w_+, w_- \in \mathbf{R}^n$, $b_+, b_- \in \mathbf{R}$. Here, each hyperplane is close to one of the two classes and is at least one distance from the other class points. The formulation of TWSVMis as follows:

$$\begin{aligned} \min_{w_+,b_+} \ & \tfrac{1}{2}\|Aw_+ + e_+b_+\|_2^2 + C_1 e_-^T \xi \\ \text{s.t.} \ & -(Bw_+ + e_-b_+) + \xi \geq e_-, \xi \geq 0 \end{aligned} \tag{2}$$

and

$$\begin{aligned} \min_{w_-,b_-} \ & \tfrac{1}{2}\|Bw_- + e_-b_-\|_2^2 + C_2 e_+^T \eta \\ \text{s.t.} \ & (Aw_+ + e_+b_+) + \eta \geq e_+, \eta \geq 0, \end{aligned} \tag{3}$$

where C_1, C_2 are positive numbers, e_+, e_- are vectors of ones of corresponding dimensions. The nonparallel hyperplanes Eq. (1) can be obtained by solving Eqs. (2) and (3). Then the new point is classified by following decision function

$$x^T w_r + b_r = \min_{r=+,-} \mid x^T w_r + b_r \mid . \tag{4}$$

2.2 Chance Constrained Twin Support Vector Machine

When uncertainties exists in the data points, the TWSVM model need to be modified to contain the uncertain information. Suppose there are l_1 and l_2 training data points in $\mathbf{R}^n$, use $\widetilde{A_i} = [A_{i1}, \cdots, A_{in}]$, $i = 1, \cdots, l_1$ to denote the uncertain data points and the label is positive $+1$. And let $\widetilde{B_i} = [B_{i1}, \cdots, B_{in}]$, $i = 1, \cdots, l_2$ to denote the uncertain data points and the label is negative -1 respectively. Then $\widetilde{A} = [\widetilde{A_1}, \cdots, \widetilde{A_{l_1}}]^T$ and $\widetilde{B} = [\widetilde{B_1}, \cdots, \widetilde{B_{l_2}}]^T$ represent two data sets. The chance-constrained program is to ensure the small probability of misclassification for the uncertain data. The chance-constrained TWSVM(CC-TWSVM) formulation is

$$\begin{aligned} \min_{w_+,b_+} \ & \tfrac{1}{2}\mathbb{E}\{\|\widetilde{A}w_+ + e_+b_+\|_2^2\} + C_1 \sum_{i=1}^{l_1} \xi_i \\ \text{s.t.} \ & \mathbb{P}\{-(\widetilde{B_i}w_+ + b_+) \leq 1 - \xi_i\} \leq \varepsilon \\ & \xi_i \geq 0, \quad i = 1, \cdots, l_1 \end{aligned} \tag{5}$$

and

$$\begin{aligned} \min_{w_-,b_-} \ & \tfrac{1}{2}\mathbb{E}\{\|\widetilde{B}w_- + e_-b_-\|_2^2\} + C_2 \sum_{i=1}^{l_2} \eta_i \\ \text{s.t.} \ & \mathbb{P}\{(\widetilde{A_i}w_- + b_-) \leq 1 - \eta_i\} \leq \varepsilon \\ & \eta_i \geq 0, \quad i = 1, \cdots, l_2. \end{aligned} \tag{6}$$

where $\mathbb{E}\{\cdot\}$ denote the expectation under corresponding distribution, C_1, C_2 are positive numbers, e_+, e_- are vectors of ones of corresponding dimensions, $0 < \varepsilon < 1$ is a parameter close to 0 and $\mathbb{P}\{\cdot\}$ is the probability distribution. The model ensures an upper bound on the misclassification probability.

3 Fuzzy Chance Constrained Twin Support Vector Machine

In this section, we propose a chance constrained twin support vector machine to process uncertain data points. When uncertainties exists in the data points, the TWSVM model need to be modified to contain the uncertain information. Like CC-TWSVM, suppose there are l_1 and l_2 training data points in $\mathbf{R}^n$, use $\widetilde{A_i} = [A_{i1}, \cdots, A_{in}]$, $i = 1, \cdots, l_1$ to denote the uncertain data points and the label is positive $+1$. And let $\widetilde{B_i} = [B_{i1}, \cdots, B_{in}]$, $i = 1, \cdots, l_2$ to denote the uncertain data points and the label is negative -1 respectively. Then $\widetilde{A} = [\widetilde{A_1}, \cdots, \widetilde{A_{l_1}}]^T$ and $\widetilde{B} = [\widetilde{B_1}, \cdots, \widetilde{B_{l_2}}]^T$ represent two data sets. By considering the contribution of different misclassified points, the chance-constrained program is to ensure the minimum misclassification rate for the uncertain data. The chance-constrained TWSVM(FCC-TWSVM) formulation is

$$\begin{aligned} \min_{w_+, b_+} \ & \tfrac{1}{2}\mathbb{E}\{\|\widetilde{A}w_+ + e_+b_+\|_2^2\} + C_1 \sum_{i=1}^{l_1} t_i\xi_i \\ \text{s.t.} \ & \mathbb{P}\{-(\widetilde{B_i}w_+ + b_+) \leq 1 - \xi_i\} \leq \varepsilon \\ & \xi_i \geq 0, \quad i = 1, \cdots, l_1 \end{aligned} \tag{7}$$

and

$$\begin{aligned} \min_{w_-, b_-} \ & \tfrac{1}{2}\mathbb{E}\{\|\widetilde{B}w_- + e_-b_-\|_2^2\} + C_2 \sum_{i=1}^{l_2} t_i\eta_i \\ \text{s.t.} \ & \mathbb{P}\{(\widetilde{A_i}w_- + b_-) \leq 1 - \eta_i\} \leq \varepsilon \\ & \eta_i \geq 0, \quad i = 1, \cdots, l_2. \end{aligned} \tag{8}$$

where $\mathbb{E}\{\cdot\}$ denote the expectation under corresponding distribution, C_1, C_2 are positive numbers, $t_i \in (1, 1]$ denotes the fuzzy membership of positive and negative samples, $0 < \varepsilon < 1$ is a parameter close to 0 and $\mathbb{P}\{\cdot\}$ is the probability distribution. The model ensures an upper bound on the misclassification probability, but two quadratic optimization problems (7) and (8) with chance constrained are typically non-convex, so the model is very hard to solve.

The work so far to deal with the chance constraint is to transfer them by different bounding inequalities. When the mean and covariance matrix of uncertain data points are known, then multivariate bound [3,16,17] by robust optimization can be used to express the chance constraints in special condition [4,21].

Lemma 1 [3,16]**.** *Let $X \sim (\mu, \Sigma)$ denote random vector X with mean μ and covariance matrix Σ, the multivariate Chebyshev inequality states that for any closed convex set S, the supremum of the probability that X take a value in S is*

$$\begin{aligned} \sup_{X \sim (\mu, \Sigma)} \ & \mathbb{P}\{X \in S\} = \tfrac{1}{1+d^2} \\ & d^2 = \inf_{X \in S} (X - \mu)^T \textstyle\sum^{-1} (X - \mu). \end{aligned} \tag{9}$$

Theorem 1. *Assume the first and second moment information of random variables $\widetilde{A_i}$ and $\widetilde{B_i}$ are known. Let $\mu_i^+ = \mathbb{E}[\widetilde{A_i}]$ and $\mu_i^- = \mathbb{E}[\widetilde{B_i}]$ be the mean* *vector seperately. And let* $\sum_i^+ = \mathbb{E}[(\widetilde{A_i} - \mu_i^+)^T(\widetilde{A_i} - \mu_i^+)]$ *and* $\sum_i^- = \mathbb{E}[(\widetilde{B_i} - \mu_i^-)^T(\widetilde{B_i} - \mu_i^-)]$ *be the covariance matrix of the two data set uncertain points respectively. Then the problems (7) and (8) could be reformulated respectively as:*

$$\begin{aligned} &\min_{w_+,b_+} \tfrac{1}{2} w_+^T G^+ w_+ + w_+^T {\mu^+}^T b_+ + \tfrac{1}{2} l_1 b_+^2 + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\ &s.t. \ \ -(\mu_i^- w_+ + b_+) \geq 1 - \xi_i + k \| {\textstyle\sum_i^{-\frac{1}{2}}} w_+ \|, \quad \xi_i \geq 0 \end{aligned} \tag{10}$$

and

$$\begin{aligned} &\min_{w_-,b_-} \tfrac{1}{2} w_-^T G^- w_- + w_-^T \mu^{-T} b_- + \tfrac{1}{2} l_2 b_-^2 + C_2 \sum_{i=1}^{l_2} t_i \eta_i \\ &s.t. \ \ \mu_i^+ w_- + b_- \geq 1 - \eta_i + k \| {\textstyle\sum_i^{+\frac{1}{2}}} w_- \|, \quad \eta_i \geq 0, \end{aligned} \tag{11}$$

where $k = \sqrt{\frac{1-\varepsilon}{\varepsilon}}$ *and*

$$G^+ = \sum_{i=1}^{l_1} ({\mu_i^+}^T \mu_i^+ + \Sigma_i^+), \quad \mu^+ = \sum_{i=1}^{l_1} \mu_i^+$$

with

$$G^- = \sum_{i=1}^{l_2} ({\mu_i^-}^T \mu_i^- + \Sigma_i^-), \quad \mu^- = \sum_{i=1}^{l_2} \mu_i^-.$$

Proof. Now we prove that problem (7) can be reformulated to Eq. (10). In fact, it follows from Eq. (7) that

$$\begin{aligned} &\min_{w_+,b_+} \frac{1}{2} \mathbb{E}\{\|\widetilde{A} w_+ + e_+ b_+\|_2^2\} + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\ &= \min_{w_+,b_+} \frac{1}{2} \mathbb{E}\{\sum_{i=1}^{l_1} (\widetilde{A_i} w_+)^2 + 2 \sum_{i=1}^{l_1} \widetilde{A_i} w_+ b_+ + l_1 b_+^2\} + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\ &= \min_{w_+,b_+} \frac{1}{2} w_+^T \mathbb{E}\{\sum_{i=1}^{l_1} \widetilde{A_i}^T \widetilde{A_i}\} w_+ + \sum_{i=1}^{l_1} \mathbb{E}\{\widetilde{A_i}\} w_+ b_+ + \frac{1}{2} l_1 b_+^2 + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\ &= \min_{w_+,b_+} \frac{1}{2} w_+^T \sum_{i=1}^{l_1} \mathbb{E}\{\widetilde{A_i}^T \widetilde{A_i}\} w_+ + \sum_{i=1}^{l_1} \mathbb{E}\{\widetilde{A_i}\} w_+ b_+ + \frac{1}{2} l_1 b_+^2 + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\ &= \min_{w_+,b_+} \frac{1}{2} w_+^T \sum_{i=1}^{l_1} (\mathbb{E}^T\{\widetilde{A_i}\} \mathbb{E}\{\widetilde{A_i}\} + \Sigma_i^+) w_+ + \sum_{i=1}^{l_1} \mathbb{E}\{\widetilde{A_i}\} w_+ b_+ + \frac{1}{2} l_1 b_+^2 + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\ &= \min_{w_+,b_+} \frac{1}{2} w_+^T \sum_{i=1}^{l_1} ({\mu_i^+}^T \mu_i^+ + \Sigma_i^+) w_+ + \sum_{i=1}^{l_1} \mu_i^+ w_+ b_+ + \frac{1}{2} l_1 b_+^2 + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\ &= \min_{w_+,b_+} \frac{1}{2} w_+^T G^+ w_+ + w_+^T \mu^T b_+ + \frac{1}{2} l_1 b_+^2 + C_1 \sum_{i=1}^{l_1} t_i \xi_i, \end{aligned}$$

where

$$G^+ = \sum_{i=1}^{l_1}(\mu_i^{+^T}\mu_i^+ + \Sigma_i^+), \quad \mu^+ = \sum_{i=1}^{l_1}\mu_i^+.$$

Moreover, for the constraint of (7), we know that the set $\{-(\widetilde{B}w_+ + e_+b_+) \leq 1-\xi\}$ is a half-space produced by a hyperplane and so it is a closed convex set. Using Lemma 1, we obtain

$$\sup_{\widetilde{B}_i\sim(\mu_i^-,\Sigma_i^-)} \mathbb{P}\{-(\widetilde{B}_i w_+ + b_+) \leq 1-\xi_i\} = \tfrac{1}{1+d^2}$$
$$d^2 = \inf_{-(Xw_+ + b_+)\leq 1-\xi_i} (X-\mu_i^-)^T \Sigma_i^{+^{-1}}(X-\mu_i^-).$$

Now we claim that $-(\mu_i^+ w_+ + b_+) > 1-\xi_i$. Suppose that $-(\mu_i^+ w_+ + b_+) \leq 1-\xi_i$. Then by taking $X = \mu_i^-$, we can get $d^2 = 0$ and

$$\sup_{\widetilde{B}_i\sim(\mu_i^-,\Sigma_i^-)} \mathbb{P}\{-(\widetilde{B}_i w_+ + b_+) \leq 1-\xi_i\} = 1,$$

which is a contradiction to $0 < \varepsilon < 1$. Let

$$u_i = \Sigma_i^{-^{-\frac{1}{2}}}(X-\mu_i), \quad v_i = -\Sigma_i^{-^{\frac{1}{2}}} w_+, \quad \gamma_i = \mu_i^+ w_+ + b_+ + 1 - \xi_i.$$

Then $\gamma_i < 0$ and

$$d^2 = \inf_{v_i^T u_i \leq \gamma_i} u_i^T u_i.$$

Consider the Lagrangian

$$L(u_i,\lambda_i) = u_i^T u_i + \lambda_i(v_i^T u_i - \gamma_i), \quad \lambda_i \geq 0.$$

By taking the derivative to be zero, we have $2u_i = \lambda_i v_i$ and $v_i^T u_i = \gamma_i$. Thus,

$$\begin{aligned} d^2 &= \inf_{-(Xw_+ + b_+)\leq 1-\xi_i} (X-\mu_i^-)^T \Sigma_i^{-^{-1}}(X-\mu_i^-) \\ &= u_i^T u_i \\ &= \frac{\gamma_i^2}{v_i^T v_i} \\ &= \frac{(\mu_i^- w_+ + b_+ - 1 + \xi_i)^2}{w^{+^T}\Sigma_i^- w^+}. \end{aligned}$$

For the constraint (7), by

$$\sup_{X\sim(\mu_i^-,\Sigma_i^-)} \mathbb{P}\{-(Xw_+ + b_+) \leq 1-\xi_i\} < \varepsilon,$$

it is easy to see that $\frac{1}{1+d^2} \leq \varepsilon$ and so $d^2 \geq \frac{1-\varepsilon}{\varepsilon}$. Since $-(\mu_i^+ w_+ + b_+) - 1 + \xi_i > 0$, one has

$$-(\mu_i^- w_+ + b_+) \geq 1 - \xi_i + k\|\textstyle\sum_i^{-^{\frac{1}{2}}} w_+\|, \quad \xi_i \geq 0.$$

Similarly, we know that problem (8) can be expressed as (11). This completes the proof. □

For any non-zero vector (w_+, b_+), we have

$$\begin{aligned}
&\frac{1}{2}w_+^T G^+ w_+ + w_+^T {\mu^+}^T b_+ + \frac{1}{2} l_1 b_+^2 \\
&= (w_+, b_+)^T \begin{bmatrix} \frac{1}{2}G^+ & \frac{1}{2}{\mu^+}^T \\ \frac{1}{2}\mu^+ & \frac{1}{2}l_1 \end{bmatrix} (w_+, b_+) \\
&= \frac{1}{2}\mathbb{E}\{\|\widetilde{A}w_+ + e_+ b_+\|_2^2\} \\
&\geq 0.
\end{aligned}$$

On the other hand, it is easy to see that the constraint of problem (10) is a form of second cone.

Let

$$H^+ = \frac{1}{2}\begin{bmatrix} G^+ & {\mu^+}^T \\ \mu^+ & l_1 \end{bmatrix}. \tag{12}$$

Then the matrix H^+ is positive semi-define. To ensure the strict convexity of problem (10), we can always append a perturbation ϵI ($\epsilon > 0$, I is the identity matrix) such that the matrix $H^+ + \epsilon I$ is positive define. Without loss of generality, suppose that H^+ is positive define.

Theorem 2. *The dual problems of chance-constrained TWSVM models (10) and (11) can be formulated as the following models*

$$\begin{aligned}
&\max_{\lambda_i, \nu} \sum_{i=1}^{l_1} \lambda_i - \tfrac{1}{2}{s_i^+}^T {H_1^+}^T G^+ H_1^+ s_i^+ - \tfrac{1}{2} l_1 {s_i^+}^T {H_2^+}^T H_2^+ s_i^+ - \mu_i^+ H_1^+ s_i^+ H_2^+ s_i^+ \\
&s.t. \ -\left(\sum_{i=1}^{l_1} \lambda_i \left({\mu_i^-}^T + k{\Sigma_i^-}^{\frac{1}{2}} \nu\right), \sum_{i=1}^{l_1} \lambda_i\right) = s_i^+ \\
&\qquad 0 \leq \lambda_i \leq C_1 t_i, \quad \|\nu\| \leq 1
\end{aligned} \tag{13}$$

and

$$\begin{aligned}
&\max_{\gamma_i, \upsilon} \sum_{i=1}^{l_2} \gamma_i - \tfrac{1}{2}{s_i^-}^T {H_1^-}^T G^- H_1^- s_i^- - \tfrac{1}{2} l_2 {s_i^-}^T {H_2^-}^T H_2^- s_i^- - \mu_i^+ H_1^+ s_i^- H_2^+ s_i^- \\
&s.t. \ -\left(\sum_{i=1}^{l_2} \gamma_i \left({\mu_i^+}^T - k{\Sigma_i^+}^{\frac{1}{2}} \upsilon\right), \sum_{i=1}^{l_2} \gamma_i\right) = s_i^- \\
&\qquad 0 \leq \gamma_i \leq C_2 t_i, \quad \|\upsilon\| \leq 1,
\end{aligned} \tag{14}$$

where

$${H^+}^{-1} = [H_1^+, H_2^+], \quad {H^-}^{-1} = [H_1^-, H_2^-].$$

Proof. We only need to prove that the dual problem of Eq. (10) can be formulated as Eq. (13). The Lagrangian is given by

$$\begin{aligned}
L(w_+, b_+, \xi, \lambda, \beta) =& \frac{1}{2}w_+^T G^+ w_+ + w_+^T {\mu^+}^T b_+ + \frac{1}{2}l_1 b_+^2 + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\
&- \sum_{i=1}^{l_1} \lambda_i \left(-(\mu_i^- w_+ + b_+) - 1 + \xi_i - k\|{\Sigma_i^-}^{\frac{1}{2}} w_+\|\right) - \sum_{i=1}^{l_1} \beta_i \xi_i,
\end{aligned} \tag{15}$$

where $\lambda_i, \beta_i \geq 0$.

Recall that for any $x \in \mathbf{R}^n$, we have the relationship $\|x\| = \max_{\|y\| \leq 1} x^T y$. Then the equivalent model of Eq. (15) are given as follows:

$$L_1(w_+, b_+, \xi, \lambda, \beta, \nu) = \frac{1}{2} w_+^T G^+ w_+ + w_+^T {\mu^+}^T b_+ + \frac{1}{2} l_1 b_+^2 + C_1 \sum_{i=1}^{l_1} t_i \xi_i \\ - \sum_{i=1}^{l_1} \lambda_i \left(-(\mu_i^- w_+ + b_+) - 1 + \xi_i - k(\Sigma_i^{-\frac{1}{2}} w_+)^T \nu \right) - \sum_{i=1}^{l_1} \beta_i \xi_i, \tag{16}$$

where $\lambda_i,\ \beta_i \geq 0$, $\|\nu\| \leq 1$. The Lagrangian L_1 has the same optimal value as L when maximized with respect to ν's subject to the constraint $\|\nu\| \leq 1$. Therefore,

$$L(w_+, b_+, \xi, \lambda, \beta) = \max_{\|\nu\| \leq 1} L_1(w_+, b_+, \xi, \lambda, \beta, \nu).$$

Similar to the discussion of Sect. 5 of [6], we know that solving Eq. (10) is equivalent to finding the saddle-point of the Lagrangian L_1. This fact combining the convexity implies

$$\begin{aligned} &\min_{w_+, b_+, \xi} \max_{\lambda_i, \beta_i} L(w_+, b_+, \xi, \lambda, \beta) \\ = &\min_{w_+, b_+, \xi} \max_{\lambda_i, \beta_i, \|\nu\| \leq 1} L_1(w_+, b_+, \xi, \lambda, \beta, \nu) \\ = &\max_{\lambda_i, \beta_i, \|\nu\| \leq 1} \min_{w_+, b_+, \xi} L_1(w_+, b_+, \xi, \lambda, \beta, \nu). \end{aligned} \tag{17}$$

By eliminating the primal variables in Eq. (17), we can obtain the dual problem. Taking partial derivatives of L_1 with respect to $w_+, b_+,$ and ξ, respectively, one has

$$\begin{cases} \frac{\partial L_1(w_+, b_+, \xi, \lambda, \beta)}{\partial w_+} = G^+ w_+ + {\mu^+}^T b_+ + \sum_{i=1}^{l_1} \lambda_i {\mu_i^-}^T + k \sum_{i=1}^{l_1} \lambda_i \Sigma_i^{-\frac{1}{2}} \nu, \\ \frac{\partial L_1(w_+, b_+, \xi, \lambda, \beta)}{\partial b_+} = w_+^T {\mu^+}^T + l_1 b_+ + \sum_{i=1}^{l_1} \lambda_i, \\ \frac{\partial L_1(w_+, b_+, \xi, \lambda, \beta)}{\partial \xi_i} = C_1 t_i - \lambda_i - \beta_i. \end{cases} \tag{18}$$

By equating the equations in Eq. (18) to zero, we can obtain following formulations

$$\begin{cases} G^+ w_+ + {\mu^+}^T b_+ + \sum_{i=1}^{l_1} \lambda_i {\mu_i^-}^T + k \sum_{i=1}^{l_1} \lambda_i \Sigma_i^{-\frac{1}{2}} \nu = 0, \\ w_+^T {\mu^+}^T + l_1 b_+ + \sum_{i=1}^{l_1} \lambda_i = 0, \\ C_1 t_i - \lambda_i - \beta_i = 0. \end{cases} \tag{19}$$

Thus, by solving Eq. (19) and considering (12), we can infer that

$$H^+[w_+, b_+] = -\left(\sum_{i=1}^{l_1} \lambda_i \left({\mu_i^-}^T + k \Sigma_i^{-\frac{1}{2}} \nu \right), \sum_{i=1}^{l_1} \lambda_i \right) = s_i^+. \tag{20}$$

Since H^+ is positive define, it is easy to see the solution of Eq. (20) can be obtained as

$$\begin{cases} w_+ = -H_1^+\left(\sum_{i=1}^{l_1}\lambda_i\left(\mu_i^{-^T} + k\Sigma_i^{-\frac{1}{2}}\nu\right), \sum_{i=1}^{l_1}\lambda_i\right) = H_1^+ s_i^+, \\ b_+ = -H_2^+\left(\sum_{i=1}^{l_1}\lambda_i\left(\mu_i^{-^T} + k\Sigma_i^{-\frac{1}{2}}\nu\right), \sum_{i=1}^{l_1}\lambda_i\right) = H_2^+ s_i^+, \end{cases} \tag{21}$$

where $H^{+^{-1}} = [H_1^+, H_2^+]$.

According to Eqs. (16), (19) and (21), we can obtain the dual problem of Eq. (10) as follows

$$\begin{aligned} &\min_{\lambda_i, u} \sum_{i=1}^{l_1}\lambda_i - \tfrac{1}{2}s_i^{+^T} H_1^{+^T} G^+ H_1^+ s_i^+ - \tfrac{1}{2} l_1 s_i^{+^T} H_2^{+^T} H_2^+ s_i^+ - \mu_i^+ H_1^+ s_i^+ H_2^+ s_i^+ \\ &\text{s.t.} -\left(\sum_{i=1}^{l_1}\lambda_i\left(\mu_i^{-^T} + k\Sigma_i^{-\frac{1}{2}}\nu\right), \sum_{i=1}^{l_1}\lambda_i\right) = s_i^+ \\ &\qquad 0 \le \lambda_i \le C_1 t_i, \quad \|\nu\| \le 1. \end{aligned}$$

Similarly, we know that the dual problem of Eq. (11) can be expressed as Eq. (14). This completes the proof. □

4 Numerical Experiments

In this section, our FCC-TWSVM model is illustrated by numerical test based on two types of data sets. The first test is implemented to certify the performance of our FCC-TWSVM by artificial data. And in second test, we also test the performance of FCC-TWSVM model on real-word classifying data sets from UCI Machine Learning Repository. All results were averaged on 10 train-test experiments and carried out by Matlab R2012a with 2.5GHz CPU, 2.5G usable RAM. The SeDuMi [1] software is employed to solve the SOCP problems of FCC-TWSVM.

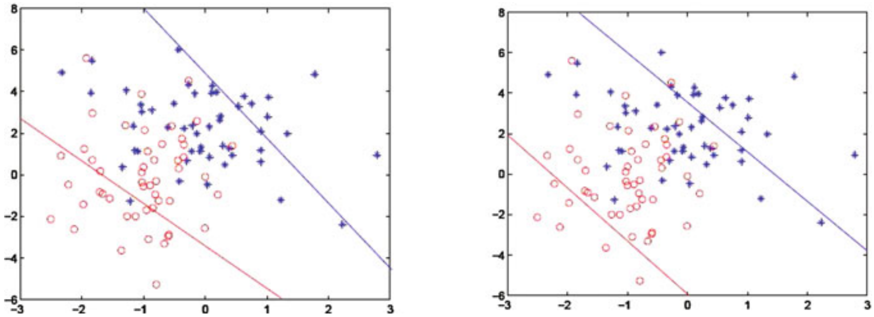
4.1 Fuzzy Membership Function

The design of fuzzy membership function is the key to the fuzzy algorithm using fuzzy technology. In this paper, we use class center method to generate fuzzy membership. Firstly, we denote the mean of class +1 as class-center x_+ and the mean of class -1 as class center x_-, respectively. The radius of each class r+ and r- are the farthest distance between the each class training points and its class-center, respectively, namely $r_+ = \max_{x_i, y_i = 1} \|x_i - x_+\|$ and $r_- = \max_{x_i, y_i = -1} \|x_i - x_-\|$. Fuzzy membership t_i is a function of the mean and radius of each class

$$t_i = \begin{cases} 1 - \frac{\|x_i - x_+\|}{r_+ + \sigma} & \text{if } \ y_i = 1 \\ 1 - \frac{\|x_i - x_-\|}{r_- + \sigma} & \text{if } \ y_i = -1, \end{cases}$$

where $\sigma > 0$ is used to avoid the case $t_i = 0$.

[1] http://sedumi.ie.lehigh.edu/.



(a) ε = 0.1 (b) ε = 0.01

Fig. 1. The performance of FCC-TWSVM in the first data set.

4.2 Artificial Data

To give an intuitive performance of FCC-TWSVM, we construct uncertain sets of 2-dimension data generated randomly from two normal distribution. The contribution of binary classes of the first data is

$$\mu^+ = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad \Sigma^+ = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad \mu^- = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \Sigma^- = \begin{bmatrix} 7 & 0 \\ 0 & 3 \end{bmatrix}.$$

Figures 1 shows that the performance of FCC-TWSVM to two uncertain data set points. In numerical experiments, different data points are generated by respective distribution. In each data set, +1 class is generated by normal distribution (μ^+, Σ^+) and -1 class is generated by normal distribution (μ^-, Σ^-). Each class has 50 points, and 20 points are randomly picked as the training points, the other points are the test points. In Fig. 1, the blue stars are the points of +1 class, while -1 class with the red circles. The blue and red lines are the separating hyperplane (2-D) that we look for. In fact, the value of parameter ε also affects the determination of two hyperplanes. Figure 1(a) and (b) perform the effect of various parameters.

4.3 Application

China's VC [7] market is now in a substantial expansion of scale, while of the decline in return. The industry system is between preliminary cooperation and in-depth cooperation. The former is reflected in the VC firms, and the latter is reflected in the investment groups.

We call such structure a community which is closely related to each other and has far more links than other nodes in social network analysis method. From the sociology level, a community is a social network with strong relationships. The preference-dependent phenomenon in random network model has been widely accepted, but the condition of new VC firm's attachment is still unknown.

The formation of communities is determined by the attributes of VC firms. We select five basic attributes: capital scale, registration place, institution type, number of investments, age. Capital scale stands for the strength of VC firms. More capital controlled means stronger bargain power, which means the cooperative VC firms could make the advantage position in the equity pricing. If capital scale plays an important role in the decision tree, it means that big capital trends to cooperate with big capital as well as small capital, where there is a polarization. Registration place can be deemed to be the place where the headquarters or the institution with high administrative authority is, at least. It shows one of the most active areas in social and economic activities. Institution type indicates the main business type of VC firms. In addition to VC, PE, there are PIPE, strategic investment, angel investment, broker direct investment, which are classified as other categories. The number of investments represents the active level of VC firms. Although there is a certain correlation between capital scale, age, and the number of investments, taking these three indicators into consideration simultaneously could reduce the impact of relevance, to a certain extent.

In practice, the distribution properties are often unknown but need to be estimated from data. For example, if an uncertain data point $\widetilde{x_i} = [\widetilde{x_{i_1}}, \cdots, \widetilde{x_{i_n}}]^T$ has N samples x_{i_k}, $k = 1, \cdots, N$, then the sample mean $\overline{x_i} = \frac{1}{N}\sum_{i=1}^{N} x_{i_k}$ is used to estimate the mean vector $\mu_i = \mathbb{E}[\widetilde{x_i}]$, and the sample covariance

$$S_i = \frac{1}{N-1}\sum_{i=1}^{N}(x_{i_k} - \overline{x_i})(x_{i_k} - \overline{x_i})^T$$

is used to estimate the covariance matrix

$$\Sigma_i = \mathbb{E}[(\widetilde{x_i} - \mu_i)(\widetilde{x_i} - \mu_i)^T].$$

However, these could cause possible estimation errors. Some special cases were proposed when the mean vector μ_i and covariance matrix Σ_i may not exactly known. Panos M. Pardalos et al. [24] has discussed the way to processing these special cases. In our practical experiments, similar to Pardalos, we employ mentioned methods to modify the estimation and make the result easier to interpret.

Since the data sets are uncertain, the performance measures are worth discussed. Ben-Tal et al. [2] proposed using nominal error and optimal error to evaluate the performance. In our experiment, we choose these index to calculate the accuracy of our model.

The expression for NomErr is

$$\text{NomErr} = \frac{\sum_i \mathbf{1}_{y_i^{pre} \neq y_i}}{\text{the amount of training data}} \times 100\%.$$

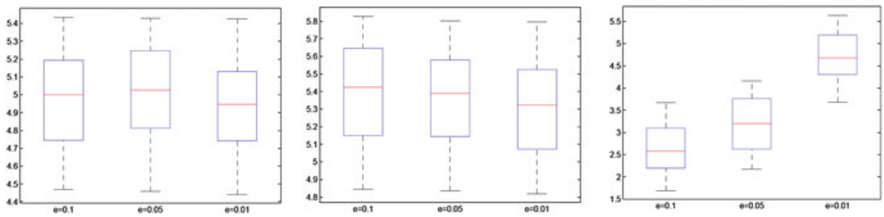
The optimal error is based on the probability of misclassification. The chance constraints in the model (7), (8) can be reformulated to Eqs. (10) and (11), then we can derive the least value of ε called ε_{opt}. Then the OptErr of data point x_i is

$$\text{OptErr} = \begin{cases} 1 & \text{if } \; y_i^{\text{pre}} \neq y_i \\ \varepsilon_{opt} & \text{if } \; y_i^{\text{pre}} = y_i. \end{cases}$$

And the OptErr of whole test set is

$$\text{OptErr} = \frac{\sum_i \text{OptErr}_i}{\text{the amount of training data}} \times 100\%.$$

We tested the from Qingke data set. Now each test data point x_i has 50 replicates x_{i_k}. The class label y_i^{pre} is decided by the majority label of the replicates based on Eq. 4. For the OptErr, the sample mean and sample covariance ε_{opt}. The real data sets have too many of the features to test the performance of FCC-TWSVM model. The results over 10 times experiments are shown in Fig. 2



(a) NomErr (b) OptErr (c) Training time

Fig. 2. The performance of FCC-TWSVM in the VC data set

The average results for VC data set and Ionosphere set are shown in Fig. 2. The boxplots of the results for two different sets show that the misclassification rate decreases when ε reduces. In addition, the OptErr is always bigger than NomErr. The experiment time is stable for different parameters ε. We can draw a conclusion that regional disparity is the biggest factor in the difference of community. There is a high information asymmetry between enterprises and VC firms, and the latter usually give more attention to local projects.

5 Conclusions

A new chance constrained twin support vector machine (FCC-TWSVM) via chance constrained programming formulation for classification was proposed, which can deal data with measurement noise efficiently. This paper studied twin support vector machine classification when data points are uncertain statistically. With some properties known for the distribution, the FCC-TWSVM model was used to ensure the small probability of misclassification for the uncertain data. The FCC-TWSVM model could be transformed to second-order cone programming (SOCP) by the properties of moment information of uncertain data and the dual problem of SOCP model was also introduced. Then we obtained the twin hyperplanes by calculating the dual problem. In addition, we also showed the performance of FCC-TWSVM model in artificial data and real data by numerical experiments. In the future work, how to further make the model more robust is under our consideration. In addition, dealing the situation of nonlinear classification with chance constrained is also interesting.

Acknowledgements. This paper is supported by the Fundamental Research Funds for the Central Universities of Sichuan University (skqy201646, skgt201602).

References

1. Ben-Tal A, Nemirovski A (2008) Selected topics in robust convex optimization. Math Progam 112(1):125–158
2. Ben-Tal A, Bhadra S et al (2011) Chance constrained uncertain classification via robust optimization. Math Program 127(1):145–173
3. Bertsimas D, Popescu I (2005) Optimal inequalities in probability theory: a convex optimization approach. SIAM J Optim 15(3):780–804
4. Bhattacharyya C, Grate LR et al (2004) Robust sparse hyperplane classifiers: application to uncertain molecular profiling data. J Comput Biol 11(6):1073–1089
5. Bi J, Zhang T (2004) Support vector classification with input data uncertainty. Procof Neural Infprocsyst 17:161–168
6. Boyd S, Vandenberghe L (2004) Convex optimization. Cambridge University Press, Cambridge
7. Cao Q, Lu Y et al (2013) The roles of bridging and bonding in social media communities. J Assoc Inf Sci Technol 64(8):1671–1681
8. Carrasco M, López J, Maldonado S (2016) A second-order cone programming formulation for nonparallel hyperplane support vector machine. Expert Syst Appl 54(C):95–104
9. Chang CC, Lin CJ (2007) Libsvm: a library for support vector machines. ACM Trans Intell Syst Technol 2(3):389–396 Article 27
10. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
11. Goldfarb D, Iyengar G (2003) Robust convex quadratically constrained programs. Math Program 97(3):495–515
12. Jayadeva Khemchandani R, Chandra S (2007) Twin support vector machines for pattern classification. IEEE Trans Pattern Anal Mach Intell 29(5):905–910
13. Lee YJ, Mangasarian OL (2001) SSVM: a smooth support vector machine for classification. Comput Optim Appl 20(1):5–22
14. Lin CF, Wang SD (2002) Fuzzy support vector machines. IEEE Trans Neural Netw 13(2):464–471
15. Mangasarian OL, Wild EW (2006) Multisurface proximal support vector machine classification via generalized eigenvalues. IEEE Trans Pattern Anal Mach Intell 28(1):69–74
16. Marshall AW, Olkin I (1960) Multivariate Chebyshev inequalities. Ann Math Stat 31(4):1001–1014
17. Nemirovski A, Shapiro A (2006) Convex approximations of chance constrained programs. SIAM J Optim 17(4):969–996
18. Peng X (2010) A v-twin support vector machine (v-TWSVM) classifier and its geometric algorithms. Inf Sci 20(180):3863–3875

19. Scholkopf B, Smola AJ (2003) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge
20. Shao YH, Deng NY (2012) A coordinate descent margin based-twin support vector machine for classification. Neural Netw Official J Int Neural Netw Soc 25(1):114–121
21. Shivaswamy PK, Bhattacharyya C, Smola AJ (2006) Second order cone programming approaches for handling missing and uncertain data. J Mach Learn Res 7(7):1283–1314
22. Trafalis T, Gilbert R (2007) Robust support vector machine for classification and computatioanla issues. Optim Meth Soft 1(22):187–198
23. Trafalis TB, Gilbert RC (2006) Robust classification and regression using support vector machines. Eur J Oper Res 173(3):893–909
24. Wang X, Fan N, Pardalos PM (2015) Robust chance-constrained support vector machines with second-order moment information. Ann Oper Res 253:1–24
25. Xanthopoulos P, Pardalos PM et al (2013) Robust data mining. Springer briefs in optimization
26. Yang B, Wang MH et al (2016) Ramp loss quadratic support vector machine for classification. Nonlinear Anal Forum 1(21):101–115

A Projection Pursuit Combined Method for PPP Risk Evaluation

Xinli Zhang(✉), Tianjin Wang, and Sihan Li

Business School, Sichuan University, Chengdu 610064, People's Republic of China
zhangxinli1231@163.com

Abstract. Risk assessment is very crucial to the building and operation of PPP (Public-Private-Partnership) project in its long-term cooperation during the long life circle. Based on the risk factors and their influences, the traditional risk assessment method of PPP project is achieved by the experts experience, Therefore, the evaluation result is subjective. Under the background of Big Data, the study proposed a projection pursuit combined method to model the PPP risk, which was called projection pursuit risk assessing model. Firstly, we use Monte Carlo simulation to get the data of the risks factors and their influence. Secondly, establish a multivariate regression function with projection pursuit method to research the relationship between the multivariable risk factors and the increment of internal return rate of PPP project. Thirdly, we offer the project risk level through the analysis of the regression results. The application in the specific project of this new method shows that the new model can accurately estimate the risk level of PPP project, and explain the trend of project risk.

Keywords: PPP (Public-Private-Partnership) · Risk assessment · Projection pursuit regression · Monte Carlo simulation

1 Introduction

In recent years, PPP (Private-Public-Partnership) pattern has been introduced to China as a new mode, aiming at attracting foreign capital to participate in infrastructure construction projects to make up for the financial lack of public sectors as well as improving the operational efficiency. Thus the mode is widely regarded as a new infrastructure project financing tool.

However, due to the long cooperation process, the wide investment scale, and the large number of participants, PPP projects have more potential and complex risk factors during the whole life cycle compared with average engineering project. Especially, risk assessment is a crucial link in the risk management process. However, currently few domestic cases have involved the implementation of the PPP project, thus the PPP risk assessment mostly depends on the expert's subjective judgment, such as the fuzzy mathematical evaluation method and decision tree method often used in related research and field [3].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_128

In recent research, s the methods for PPP assessment are mainly quantitative assessment methods. For example, fuzzy-AHP-based risk assessment and fuzzy synthetic evaluation approach to PPP projects were established to assess risk factors in the PPP expressway project in China [8]. Another research identify the risk factors in PPP projects through a comprehensive literature review, and then introduce fuzzy logic in the pairwise comparison [12]. A few studies deal with uncertainty by presenting several estimates based on different values for the exogenous inputs using sensitivity analysis. Sensitivity analysis is a typical way to address uncertainty via variations in key inputs. It hence provides an insight into what happens if some variables values differ from the basic case [7].

As research shows, system dynamics (SD) modelling aims at developing a dynamic model to assess demand risk by evaluating how different variables jointly affect demand for services provided by PPP infrastructure projects. The research objectives thus involve identifying, understanding, mapping and measuring these variables and their complex interrelations [1].

FSE is a branch of fuzzy set theory (FST), a number of researchers have attempted to exploit FSE in the construction project risk management field. Yeung et al. present a model to assess the level of risk of PPP toll road projects in which experience knowledge based on linguistic variables were incorporated into the analysis using FSE [2].

In summary, the current PPP project risk assessment has a series of complex problems, for example, the original project data involves various risk factors which share no obvious function relationship between these factors. It is difficult to precisely describe the PPP risks with accurate mathematical model since most risk assessment methods are limited to subjective factors. The risk assessment is a high-dimensional problem, which the current assessing method is unable to solve. Thus the good risk management needs to establish a risk estimation model which can convert high-dimensional problem into low-dimensional problem.

Considering the deficiencies of existing evaluation methods, we present the projection pursuit regression method (PPR) based on the data analysis as an exploring method of risk assessment. The method can convert the high-dimensional risk variables into low-dimensional projection variables, meanwhile building up the relationship between the projection variables and the economic evaluation index variables [13].

The realization of the new method is further elaborated in four parts. Section 2 introduces a PPP risk assessment method with PPR; Sect. 3 offers the building and parameter optimization process of the new PPP risk assessing model, followed by a case study applying the model into assessing the risks of an ecological landscape PPP project in Sect. 4. We present our conclusions in Sect. 5.

2 Projection Pursuit Risk Assessing Method

2.1 Traditional PPP Risk Assessment Procedure

Among the existing PPP risks assessing methods, the fuzzy synthetic evaluation (FSE) method is the most prevalent, which is capable of effectively dealing with the imprecise non-numerical (linguistic) terms and accounting for the fuzziness in expert knowledge that typifies risk assessment [11]. As Fig. 1 shows, the hierarchical analysis approach utilized for assessing the risk level of principal risk factors.

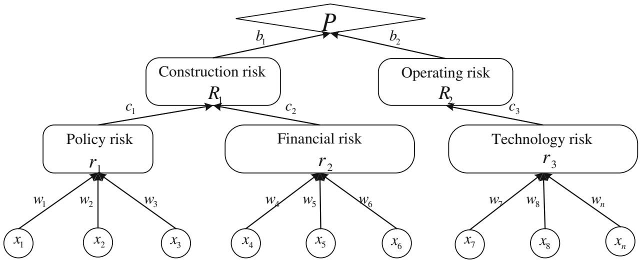


Fig. 1. FSE assessing procedure

where x_i represents the ith risk factor of the risk factor layer $(i = 1, 2 \cdots, n)$, r_i represents the sub factor layer, r_1, r_2, r_3 respectively represents the second-class risk factor layer, R_1,R_2 respectively represents the first-class risk factor layer. P represents the comprehensive risk membership of the project. w_i $(i = 1, 2, \cdots, n)$ respectively represents the i^{th} risk factor's weight in the second-class risk factor layer. c_1,c_2,c_3 respectively represents the second-class risk factor's weight in the first-class factor layer. b_1, b_2 respectively represents the first-class factor's weight in the comprehensive risk membership of the project. The relationships among these parameters can be presented as

$$X = \{x_1, x_2, x_3, \cdots, x_n\}, \tag{1}$$
$$R' = \{r_1, r_2, r_3\}, \tag{2}$$
$$R = \{R_1, R_2\}, \tag{3}$$
$$C = \{c_1, c_2, c_3\}, \tag{4}$$
$$B = \{b_1, b_2\}, \tag{5}$$
$$W = \{w_1, w_2, \cdots, w_n\}. \tag{6}$$

And then,

$$R' = W \times X, \tag{7}$$
$$R = R' \times C, \tag{8}$$
$$P = R \times B. \tag{9}$$

After several transformations of matrices, we get the value of the evaluation index. Though the FSE method has been widely used in the risk assessing field, it still has its limitations:

(1) The mechanism of how a risk factors exerts impact on the project objectives is unclear, so the actual process might be fuzzier.

(2) The interaction of various risk factors is ignored.
(3) Difficulty in accurately determining the membership.

Thus we need to discover a new method to optimize the traditional PPP assessing methodology.

2.2 Projection Pursuit Regression

Based on the analysis above, we need to find a method which can not only simplify the PPP risk assessment process compared with the traditional risk assessment process, but also improve the accuracy of risk assessment, to fulfill the purpose of redressing the defects of subjective evaluation. Thus the projection pursuit regression is introduced.

The projection pursuit regression is based on a regression method derived from projection pursuit thought [10]. Compared with the traditional methods focusing on exploring the relationship between the independent and dependent variables, this method projects multidimensional independent variables and determines the projection variables after obtaining a low dimensional projection of variables. The projection can reduce the number of variables which are used to establish the regression model, meanwhile transform the regression relationship between multiple variables and single variables into one solely between the single variables.

As for the PPP project risk assessment, it can be considered as a multi-input and single-output variable prediction problem. The ideology is, under different risk situations, we have data of set of risk values of 1 groups X_k $(k = 1, 2, \cdots, l)$ and these corresponding economic evaluation results(Y_k), then we can set up the relation between X_k and Y_k. Since X_k is a high-dimensional set of variables, we need to project these variables to low-dimensional space and find proper projecting directions z_j $(j = 1, 2, \cdots, m)$. After each projection, we'll get its corresponding residual value $\Delta y_j (\Delta y_j = y_{j'} - y_j$ where y_i is the actual value $y_{j'}$ is the fitting value) until the residual sum of squares meet certain accuracy requirement, then the parameter $j(j = 1, 2, \cdots, m)$ is optimal number of projection direction. And by minimizing a projection, finding out the projection can reflect the structural characteristics of high-dimensional data. The projection values of each direction are fitting by one-dimensional function. Then the regression function is approximated by one-dimensional fitting function [4]. The procedure of the projection pursuit method of risk assessment is shown in Fig. 2. The non-parametric regression $f(x)$ can be summed up as:

$$Y = \sum_{j=1}^{m} g_{(j)}(\sum_{i=1}^{n} \alpha_{ji}^T X), \tag{10}$$

$$\sum_{i=1}^{n} \alpha_{ji}^2 = 1, \tag{11}$$

where, X is n-dimensional independent variable x_i, Y is the response variable, g_j represents the j^{th} smooth ridge function. m is the number of approximation function, a_{ji} is the i^{th} component projection direction of the j^{th} smooth ridge function. The obtaining of big data needed in the function and algorithm of the calculation of PPR will be presented in Sect. 3.

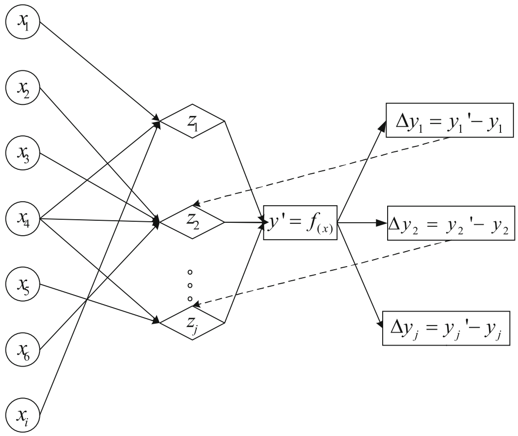


Fig. 2. The procedure of the projection pursuit method of risk assessment

3 PPP Risk Assessing Model Process and Its Parameter Optimization

3.1 The PPP Risk Assessing Model

Based on the analysis of the previous chapter, we are able to establish the relationship between the independent variable (risk variables) and the dependent variable (economic evaluation index) by the projection pursuit regression method. For every PPP project, multiple groups of risk values are not readily available to obtain. Instead they call for establishing and calculating after a large amount of data conversion and multiple model procedures. After that, the projection pursuit regression models can be summed up by Fig. 3 whose left column refers to the methods used to obtain the data, the middle column representing the specific form of the relevant data in the model, and the right column indicating the main mathematical model in the processes of data input and output.

In Fig. 3, I_k^i $(i = 1, 2 \cdots, n, \quad k = 1, 2 \cdots, s)$ represent the i^{th} risk's impact value according to s groups of historical data. $\sigma_{(I_i)}$, $\mu_{(I_i)}$ refer to the standard deviation and the mean value of the calculated values of the I_k^i. D_i represents the distribution of the i^{th} risk, I_k^i, $\mathrm{EEI}_{k'}$ respectively represents the $k'^{th}(k' = 1, 2 \cdots, l)$. Group of data of the project risk value and project economic evaluation index. α_j $(j = 1, 2, \cdots, m)$ represents the parameter in the j^{th} projection direction, $RE_{k'}$ $(k' = 1, 2 \cdots, l)$ refers to residual error. We can divide the model construction procedure into four specific steps as follows:

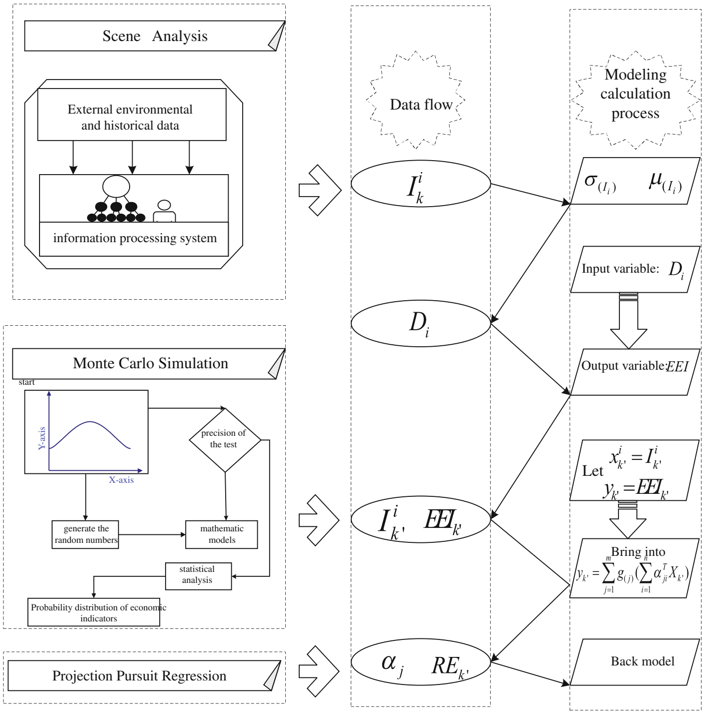


Fig. 3. The PPR risk assessing model for PPP project

Step 1. Determine the independent variables of the statistics, and fit distributions From historical data, we can obtain numerous data of different risk's impact value I_k^i of different tasks. After calculation, we can get the average value μ_i and standard deviation σ_i of the i^{th} risk's statistic value. Meanwhile, combining with experts' judgement, we can get the distributions of various risks.

$$\mu_i = \frac{\sum_{j=1}^{s} p_k^i I_k^i}{s}, \tag{12}$$

$$\sigma_i = \sqrt{\sum_{k=1}^{s} (p_k^i - \mu_i)^2 \times I_k^i}. \tag{13}$$

Step 2. Extract random number through Monte Carlo simulation and get the sample observation value.

We establish the project's economic evaluation model, and set the known variable as constant parameters variable of fixed value model. We set I_i as input random variable, which is followed by selecting economic evaluation index (EEI) according to the demand of project economic evaluation and setting it as an output variable. In order to obtain more groups of sample data, we extract random number through Monte Carlo Simulation and transform the probability distribution of random number as sampling values of input variables while repeating the experiment to meet a certain level of accuracy. At this time, we set the simulation iteration number is set as l, then we get l groups of input variables I_k^i and output variables $EEI_{k'}$ of the sample data. Now we let $x_{k'}^i = I_{k'}^i, y_{k'} = EEI_{k'}$, and establish a functional relationship between $y_{k'}$ and $x_{k'}^i$.

Step 3. Take the sample data to the projection pursuit regression model

$$y_{k'} = \sum_{j=1}^{m} g_{(j)} \left(\sum_{i=1}^{n} \alpha_{ji}^T X_{k'} \right), \tag{14}$$

$$X_{k'} = \left\{ x_{k'}^1, x_{k'}^2, \cdots, x_{k'}^i \right\} (k' = 1, 2, \cdots, l) \tag{15}$$

We send l groups of data to the PPR model, where $X_{k'}$ represents the k'^{th} set of the input variable $x_{k'}^i$ $(i = 1, 2, \cdots, n)$, $y_{k'}$ is the corresponding output variable. Other parameters mentioned in the Eq. (15) can refer to Sect. 2. Meanwhile, There are many forms to fit a variety of cell functions, such as the numerical function [13] and the polynomial function [11], a number of numerical functions, etc.

Step 4. Output the results

After the algorithm realization process, we get α_j and $RE_{k'}$ of the regression model. Generally, the smaller the $RE_{k'}$, the better the fitting effect. Then according to the distribution of *EEI*, we can arrive at some conclusions about the project's risk rank.

3.2 The Algorithm Realization Process

Since the numerical function involves an enormous numerical tables, it is not convenient to predict during the interpolation and extension [5]. Thus the paper uses multinomial to fit the unit function.

Take l groups of risk variables as input variables, and select economic evaluation index as output variables brought into Eq. (14) to determine the optimal output variables. Then optimize the number of unit functions and its corresponding projection direction parameters. In the model, the genetic algorithm is adopted to optimize the parameters as follows:

(1) Generate m projection directions randomly in the range of $[-1, 1]$, which is followed by selection, crossover and mutation of genetic algorithm to generate $3 \times n$ new projection directions [14].

(2) Respectively project m construction parameters of l sample projection on the projection direction, obtaining l projection variables for each projection direction.

(3) The polynomial is used to fit the corresponding relation between projection variables and the internal rate of return, resulting in $3 \times n$ fitting polynomials, then we can evaluate each fitting results according to the correlation coefficient d calculated through Eq. (16).

$$d = 1 - \sqrt{\frac{\sum (y' - \overline{y})^2}{\sum (y - \overline{y})^2}} (0 \leq d \leq 1). \tag{16}$$

In the Eq. (16), represents the actual value of the economical evaluation index. y' is the corresponding fitting value, and $\overline{y}$ stands for the average value of actual internal rate of return.

(4) In the $3 \times n$ projection direction, the coefficient d is calculated, and n projection directions with superior deterministic coefficient are chosen to enter the optimization process in the next round of genetic algorithm.

(5) After several times of optimization, when the deterministic coefficient's absolute value of difference between two calculations from start to finish come out less than some arbitrary positive number, the optimization of first unit function finishes.

(6) Output polynomial and its corresponding projection direction m, which lead to calculate the relative error between the fitting value and the actual value. According to the risk assessment, if the relative error is less than 10%, the model parameters are determined [6]. Otherwise, it is indicated that an arbitrary unit function can not meet the fitting requirements, and we should move to the next step.

(7) In the this step, we follow the same principle, which infers that the residual $\Delta y(y - y')$ completed in the fitting of the first step is to replace y to repeat the first step's work. We start the optimization aimed at second element function, until we have obtained fitting unit function and projection direction requirements which meet the requirements [9].

4 Case Study

4.1 Data Origin

The PPP risk assessment model based on projection pursuit regression method was validated through a case study, an ecological landscape PPP project case from Sichuan province. In this case, we identified 10 risk factors basing on the historical data, which respectively refers to cost overruns risk, construction duration extended risk, design risk, construction safety risk, operation risk, quality risk, policy changing risk, force majeure risk, tax increasing risk, and laws risk. In the PPP risk assessing model, we select (the impact value of the i^{th} risk) as independent variable, meanwhile ΔIRR as dependent variable, which can represent the increment of the internal rate of return, thus reflecting the economic viability of the project.

4.2 Input and Output Variables

We define xi as input variable, y as output variable. The distributions of each risk's impact value above can be seen in the Figs. 4 and 5, which are determined by numerous historical data Then we send the input variables and output variables into the Monte Carlo simulation model, resulting in 461 groups of sample data. The input and output variables' distributions are respectively shown in Figs. 4, 5 and 6.

4.3 Forcasting Results

With the 461 groups of sample data simulated above (436 groups are used for fitting, and 30 groups are used for testing), we can build up PPR. After experiencing algorithm realization process, the model's parameters are achieved. We can get two figures reflecting the fitting results. Figures 7 and 8 respectively shows the comparing results between the actual values and simulation values, the comparing results between the actual values and tested values, Figs. 9 and 10 respectively shows the percentage of residual error of the simulation variables and tested variables.

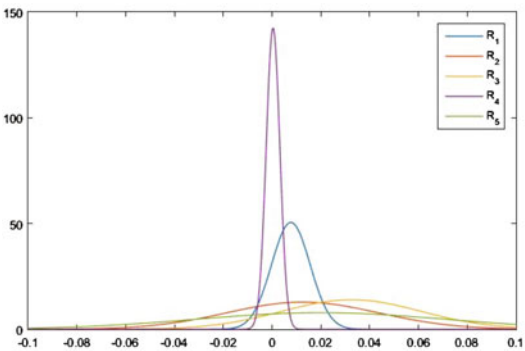


Fig. 4. The distribution of the first five risks

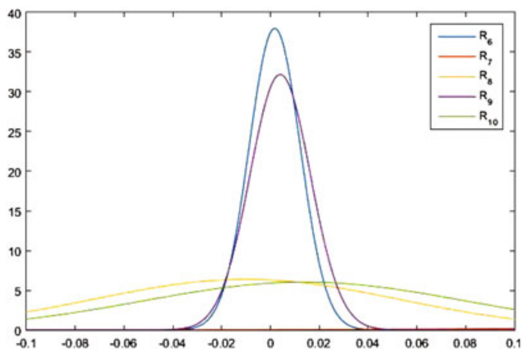


Fig. 5. The distribution of the post five risks

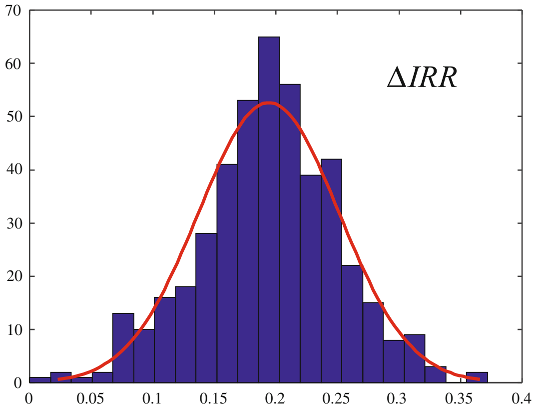


Fig. 6. The distribution of ΔIRR

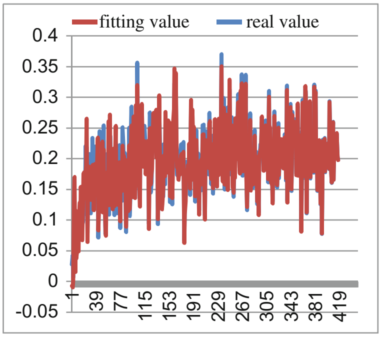


Fig. 7. The comparing results between the actual values and fitting values

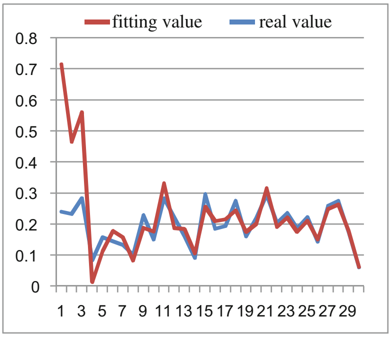


Fig. 8. The comparing results between the actual values and tested values

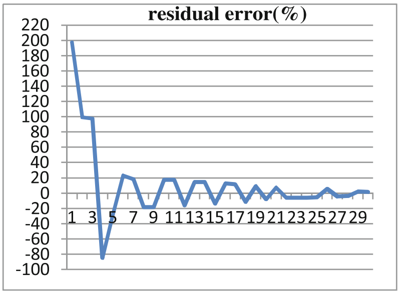


Fig. 9. The percentage of residual error of the fitting variables

Fig. 10. The percentage of residual error of the tested variables

According to the figures above, we can calculate some performance indexes in Tables 1 and 2.

Table 1. The fitting performance of PPR

The percentage of relative error	The sample size	The percentage of all samples
More than 20%	25	6%

Table 2. The tested performance of PPR

The percentage of relative error	The sample size	The percentage of all samples
More than 20%	6	20%

The results above show that PPR can get relatively accurate fitting performance for 88% sample. What's more, the majority of the relative errors are less than 10%, which indicates the variation range of the IRR is small. Namely, the anti-risk of this PPP project is good. Meanwhile, considering the risks, the investors can compare the IRR with the IRR of the benchmark yield to decide whether to invest the project.

5 Conclusion

This study introduces the projection pursuit regression method as a new risk assessment method, establishing a new risk assessment model related to the traditional risk assessment methods: Firstly we fit out the probability distribution of risk value based on historical data and expert's subjective opinions. Secondly we take the random numbers through the Monte Carlo simulation as input variables, and other economic indicators as output variables to obtain the sample data. Thirdly, we set up the interrelation between the two variables based on the data above. And then we establish a risk assessment model by projection pursuit regression method combined with the actual case, which is used to deal with the sample data regression to achieve a reasonable calculation effect. The results above shows that the model can reveal the change regulation of risk factors affecting PPP project economic evaluation index. The new assessment model can catch the risk features of PPP project influenced by many factors, to give an integrated results through dimension reduction.

The proposal of the evaluation model establishes direct interrelations between the risk factors and economic evaluation of PPP projects, reducing the intricacies of traditional risk assessment and making the risk assessment process more effective and rapid. At the same time, the evaluation model is also more accurate compared to its predecessors. However, some shortcomings also exist in this study: (1) the existing data is not sufficient enough to completely fit the probability distributions of all risks, which implies that the existing assessment of the risk still partly relies on expert assessment; (2) the evaluation model in normal circumstances is relatively stable, but part of the deviation still persists. Thus the accuracy of the algorithm calls to be improved in the optimization model in the future to improve the accuracy of risk assessment.

Acknowledgements. This paper is founded by Sichuan University (skqy201652).

References

1. Alasad R, Motawa I (2015) Dynamic demand risk assessment for toll road projects. Constr Manag Econ 33:1–19
2. Ameyaw EE, Chan APC (2015) Evaluation and ranking of risk factors in public-private partnership water supply projects in developing countries using fuzzy synthetic evaluation approach. Expert Syst Appl 42(12):5102–5116
3. Chan APC, Yeung JFY et al (2011) Empirical study of risk assessment and allocation of public-private partnership projects in China. J Manag Eng 27(3):136–148
4. Du H, Wang J et al (2008) Prediction of retention times of peptides in rplc by using radial basis function neural networks and projection pursuit regression. Chemometr Intell Lab Syst 92(1):92–99
5. Durocher M, Chebana F, Ouarda TBMJ (2016) Delineation of homogenous regions using hydrological variables predicted by projection pursuit regression. Hydrol Earth Syst Sci Discuss 20(12):4717–4729
6. Friedman JH, Stuetzle W (1981) Projection pursuit regression. J Am Stat Assoc 76(376):817–823

7. Kriger D, Shiu S, Naylor S (2006) Estimating toll road demand and revenue. NCHRP Synthesis of Highway Practice
8. Li J, Zou PXW (2011) Fuzzy AHP-based risk assessment methodology for PPP projects. J Constr Eng Manag 137(12):1205–1209
9. Mas A, Ruymgaart F (2015) High-dimensional principal projections. Complex Anal Oper Theor 9(1):35–63
10. Samarov AM (1993) Exploring regression structure using nonparametric functional estimation. J Am Stat Assoc 88(423):836–847
11. Taheriyoun M, Karamouz M, Baghvand A (2010) Development of an entropy-based fuzzy eutrophication index for reservoir water quality evaluation. Iran J Environ Health Sci Eng 7(1):1–14
12. Valipour A, Yahaya N et al (2015) A fuzzy analytic network process method for risk prioritization in freeway PPP projects: an Iranian case study. J Civ Eng Manag 21(7):933–947
13. Yamout G, Jamali D (2007) A critical assessment of a proposed public private partnership (PPP) for the management of water services in lebanon. Water Resour Manag 21(3):611–634
14. Zhang X, Wang S, Ding J (2002) Application of projection pursuit regression in estimating building's material consumption. J Sichuan Univ 29:17–23

Optimal Ownership Pattern to Control Agency Conflict in Manufacturing Industry of Pakistan

Muhammad Kaleem Khan[1], He Ying[1], Umair Akram[1(✉)], Muhammad Hashim[2], Xiaoyue Yuan[1], and Lv Gaoyu[1]

[1] School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
leo_umairinfo@yahoo.com

[2] Faculty of Management Sciences, National Textile University, Faisalabad 37610, Pakistan

Abstract. Asymmetric information always creates conflict of interest between stakeholders. Purpose of this study is to discuss agency conflict from perspective of financiers, and suggesting its remedies as controlling mechanism. Monitoring role of different types of ownerships is studied. Study has investigated the impact of four different types of ownership structures namely; managerial ownership, family ownership, institutional ownership and ownership of block-holders on agency issue. Role of financial institutions is also discussed. To explore the agency problem from view point of both equity-holders and debt-holders, Agency cost is divided into agency cost of equity (ACE) and agency cost of debt (ACD). Study has also given empirical evidences to check the robustness of each monitoring mechanism. Using financial data of top capitalized 100 companies of Pakistan from 2010 to 2015; results are drawn by fix and random effect (GLS) model. Ownership of insiders, block-holders and institutions were found to be most effective monitoring devices for firms' growth. Study also contains recommendations for investors and debtors.

Keywords: Monitoring mechanism · Agency cost · Ownership structures · Manufacturing industry

1 Introduction

After the research stream on agency cost developed by Jensen [22] extensive empirical results can be found on evidence claiming that presence of agency conflicts has an adverse effect on business sustainability. Two types of agency cost are found: Agency cost of equity (ACE) and agency cost of debt (ACD). If alignment of interest of owners and managers is missing, this is hazardous

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_129

for the company by causing agency cost of equity. If debt-holders' interest is overlooked (due to conflict of interest between managers/shareholders and debt-holder), then agency problem of debt is created. To monitor the agent's action, incurs cost that is called the agency cost.

ACE arises when principal hires agent and delegates his authority. Problem originates when agent could not meet up with principal's expectations and don't utilize firm's resources efficiently to maximize shareholders' interest. ACD is created when more risk is transferred towards debt-holder. If capital structure has higher proportion of debt, managers have strong incentives to invest on the assertion of higher return in the extremely risky projects. If the project is successful, managers and shareholders will get benefit; if not, cost will be bear by debt-holders. Managers may expropriate wealth from debt-holders in order to increase shareholders wealth or increase their own luxuries. Similarly managers can expropriate wealth from debt-holders for shareholders by following actions: getting additional debt in the case where new issued debt has equal or higher claim than the previous debt (existing debt-holders hurt), Engaging in asset substitution (switch from relatively low risky assets to highly risky assets) and by underinvestment (managers may reject profitable projects if debt-holders get benefit from such projects) and by overinvestment (when success only benefits to shareholders and failure leads to financial distress).

Present study is focusing how both types of agency cost can be reduced by decisions regarding ownership structures. Present study is addressing four different ownership structures that are most common in Pakistan: family ownership, managerial ownership, institutional ownership and ownership of block-holders. Convergence of interest hypothesis (presented by Jensen) is also going to be checked in Pakistan through the ownership of managers and family (insiders). We will be also in a position to verify whether block-holders and institutions employ their greater monitoring ability in Pakistan.

Financial institutions may alleviate not only managers' moral hazards but also agency problems between managers and creditors through a grading function ex ante and a monitoring role for their firms ex post [28]. Banks grant loan only if they are satisfied with management performance and firm's financial strength. Concept of institutional debt financing has been driven from theory of agency cost [21]. Thus institutional debt financing is indirectly involved in reducing the agency costs due to efficient monitoring of firms' decisions. Banks monitor the activities of the firms to fulfil loans extension criteria. Monitoring forces the firm to utilize assets efficiently to repay loans. Present study is focusing to check the monitoring capability of financial institutions in case of Pakistan.

Literature does not have much more about agency problem in Pakistan, making our study a quite new topic for literature. Moreover, agency issues are particularly severe issues in emerging economies [20]. Secondly ownership structure is still a new topic for researchers and there is enough room to work on this topic. So combination of these two is expected to be very valuable for Pakistan and other developing countries. We hope that our working will prove a valuable initiative for researchers of our region. Our findings and recommendations will help Pakistani manufacturing sector to enhance its performance by identifying and reducing both types of agency problem.

2 Overview of Previous Studies

Literature does not answer that which type of ownership or set of different ownerships can be optimal ownership structure. It is because researchers have not included all equity ownership styles in their study. This section compiles prominent researches regarding agency cost with different ownership structure. However, some researchers have also concluded that agency cost is irrelevant of any type of ownership structure [11,29].

Convergence of interest hypothesis advocates that when ownership and management is in the same hands, then owner-agent interests are converged [21]. Agency conflict arises due to difference of interest between owners and agents and inability of the owners to monitor agents' action. If managers become part of the ownership then this might reduce agency cost [6]. These studies also report that there is better asset utilization with managerial ownership in the firms [32,44]. Proportion of insider ownership was found very low in firms with high agency cost [2,31,46]. Leverage, dividends, managerial ownership, the percentage of outside directors and executive compensation are effective in controlling the agency cost of insider trading [5]. But there is also reporting of contradictory results of convergence of interest hypothesis [9,50]. Managerial ownership cannot reduce the agency costs of ownership within family owned firms. Private ownership sets the firm free from external governance and capital market monitoring [42]. But majority of researchers are convincing that managerial ownership acts as an effective device in controlling agency cost. For example see [1,2,4–6,13,19,24,30–32,36,38,39,44,49].

Institutional ownership is considered to be an effective monitoring mechanism due to many reasons. Qualified analysts work for institutions to analyse a firm and its financial performance [13]. Institutions also have blocks of shares that strengthen its power and can control managers efficiently to work for the attainment of financial goals [48]. Institutional investors are found successful in reducing managerial overspending in the case of firms with low growth opportunities [49], family and institutional ownerships are found to be complementary devices in monitoring the CEOs [46]. There are also findings of some studies, which don't confirm the effective monitoring role of institutions to reduce agency problem [24,36]. As previous studies found significant evidence that institutional ownership reduce agency cost [1,4,5,13,41,46,49], our research also theorizes that there is inverse relationship between institutional ownership and agency cost. Study finds this argument more convincing that institutions have greater monitoring abilities and strong influential power on the managers to direct them to do actions that increase owners' wealth. Although some studies found no influential monitoring power of institutions in reducing ACD [24,36] but this result might be due to specific environment of their countries.

Almost all the previous studies are advocating the influential monitoring power of family ownership to overcome agency conflict [6,22]. Family ownership is more committed towards stability and continued existence of the firm due to their reputations [10]. Family firms perform better than nonfamily firms [4]. Due to higher proportion of insider ownership, founding family firms have lowest agency cost [41]. Founding family ownership is also associated with a significantly lower cost of debt financing [4]. Family firms control agency problem by delivering more amount of dividends to shareholders and by employing higher amount of debt in capital structure [43]. But in family owned firms, managers act for the controlling family, but not for shareholders in general [37]. Following [3,4,41], study also hypothesizes the inverse relationship between agency cost and family ownership.

Managers of a firm, whose ownership is held by blocks of shares, are less involved in utilization of firm resources for their discretionary activities [9,14,17]. More the outsider block holding of the firm, more is the monitoring of the managerial activities [44]. Recent research has proposed a signaling theory related with block-holding ownership, according to which an entrepreneur chooses to attract an outside blockholder in order to signal his low “propensity to expropriate [45]. Large board and small non-management block-holder ownership face severe agency problems and poor corporate governance [30]. But there is also evidence that high level of block holding is responsible for higher level of agency cost in firms [41]. If ownership is intense, it can also create agency conflict [16]. Some researchers found that block-holders are not expert enough in monitoring [24].

Bank debt is an effective device to control the agency problem. Bank credit shows the credit worthiness of borrower, reducing the information asymmetry, which affects ACD [19]. Agency cost decreases with the increase in the monitoring by the banks [6]. Bank debt reduces cash holdings [17]. Institutions monitor and evaluate the company regularly for extending credit [22].

3 Empirical Implementation

In this section, we present model used to examine the impact of ownership structures on agency cost of equity and agency cost of debt. We also describe the data used and firm level variables used in this study.

$$\begin{aligned} \mathrm{ACE}_{it} = & \partial_i + \beta_1 MO_{it} + \beta_2 FO_{it} + \beta_3 IO_{it} + \beta_4 BO_{it} + \beta_5 D_{it} \\ & + \beta_6 \mathrm{DIV}_{it} + \beta_7 \mathrm{Size}_{it} + \beta_9 Q_{it} + \varepsilon_{it}, \end{aligned} \tag{1}$$

where ACE_{it} is Agency cost of equity measured in terms of assets turnover ratio: MO_{it} is managerial ownership: FO_{it} is family ownership, IO_{it} is institutional ownership, BO_{it} is ownership of block-holders: D_{it} is debt ratio: DIV_{it} is dividend payout ratio: Size_{it} is size of firm: Q_{it} represents growth opportunities for the firms. Asset utilization ratio determines management aptitude to grasp growth opportunities by efficiently employing firm resources. If this ratio is significantly good as compared to industry average, it means that high level of sales and cash flows are generated from firm assets. Whereas a low ratio of turnover represents that the managers are using assets for unproductive purpose rather than utilizing them in the activities that generate cash flows [27]. However firm with low assets turnover ratio bears more agency conflicts relative to the firm with high turnover ratio. See also [6,19,36,44]. In addition, we also assess the effect of different types of ownership structures on agency cost of debt. This notion is examined substituting the agency cost of equity variable ACE by Agency cost of debt ACD variable. The new model is described in Eq. (2). Debt agency cost is defined as conflict between shareholders and bondholders [13]. Manso took the ACD as the difference between the total value of the all-equity and levered firms [34]. We obtained the proxy of agency cost of debt (ACD) from previous studies [13,26,40]. If major proportion of capital is invested in fixed assets, then managers have no more liquid resources to use it on their own luxuries. If firms has more liquid assets, managers may found it very easy to use these assets for their own requisites or transfer cash to shareholders in form of dividends, rather than keeping it for future payments of interest payments or principal amount to debt holders [25].

$$\begin{aligned}\mathrm{ACD}_{it} = {} & \partial_i + \beta_1 MO_{it} + \beta_2 FO_{it} + \beta_3 IO_{it} + \beta_4 BO_{it} + \beta_5 D_{it} \\ & + \beta_6 \mathrm{DIV}_{it} + \beta_7 \mathrm{Size}_{it} + \beta_9 Q_{it} + \varepsilon_{it}.\end{aligned} \tag{2}$$

Inclusion of variables of ownerships of management, family, institutions and block-holders in the model is to assess the optimal ownership pattern. Addition of debt variable in the model is to test Jensen's proposition of "debt as the remedy" of agency problem [22]. In case of agency cost of equity, when firm employs more debt, then it improves the efficiency of the firm to utilize assets to produce more sales and minimize its cost [18]. It is because mangers have responsibility to pay out installment and premium regularly. If they don't do so, then creditors may take firm to court to file bankruptcy. Mangers may lose their jobs in such scenarios. So debt disciplines the mangers. In case of ACD, managers are constrained to pay fix installments against debt. This constraint directs the managers to act in such a way that they don't waste firm resources. Debt serves as a monitoring device to mitigate agency problem between owner and principal [38,44]. Firms with more debt tend to have lower agency costs [9,12,36]. Family firms take up higher debt levels to resolve agency problem [43]. Firm's capital structure affects agency cost [47]. Some studies did not find relationship of debt with agency problem [23]. The amount of free cash flow was found increased with increased level of debt [17,35,49]. Addition of Dividend payout ratio in the model is to test whether Large incremental disbursements mitigate the agency problem associated with excess funds or not [33]. Money that is not given back to outside shareholders in the form of dividend that money can be invested in negative NPV projects or can be used for the perks of insiders [7,15]. But Byrd [9] found that Dividend is not influential in defining agency cost. ACD is more influential on dividend policy as compared to that of agency cost of equity [8]. Size variable is added in model to test the notion: "Greater the size of firm, more agency cost is expected" [17,44]. Larger firms bears higher agency costs because of greater complexity in their operations [13]. That's why present study also hypothesize that there is positive relationship between firm size and agency cost. Model also has variable of growth opportunities because literature directs that if there are growth opportunities for the firms, then more cash generation in the future is expected. When firms expect incremental growth opportunities, they hold extra cash in hands to invest in future projects. It leads towards the extra cash available to the manager to be used for wasteful expenditures. Thus chances of agency problem increases with increased level of growth opportunities [17,19]. Conflicts of interest between shareholders and managers vary with growth opportunities and free cash flow [49] (Fig. 1).

Out of 819 listed companies in Karachi Stock Exchange, which is largest stock exchange of Pakistan, we selected 100 top capitalized firms. Our sample contains only non financial firms. So we ignored 353 firms related with financial services in Pakistan. By this, study got 22% representation of all the non financial firms in Pakistan, present during the period of 2010 to 2015. Our sample got representation of all the manufacturing sectors listed in KSE for the stated time. Manufacturing sector is the largest sector of the Pakistan economy.

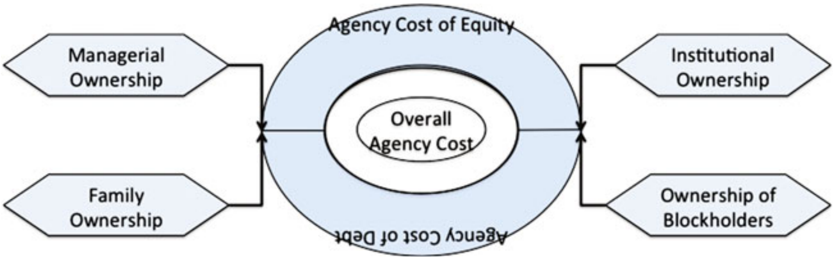


Fig. 1. Effect of ownership pattern on agency cost

Table 1. Descriptive statistics of variables

Variables	N	Minimum	Maximum	Mean	Standard dev
ATR	500	0	5.849	1.31676	1.005924
ACD	500	0.005	0.97	0.54933	0.232779
MNG	500	0	0.73	0.02428	0.078006
FML	500	0	0.921	0.15391	0.225166
INST	500	0.008	1	0.59789	0.309082
GEN	500	0	0.831	0.21549	0.195952
BLOCK	500	0	1	0.51472	0.302945
IDF	500	0	0.926	0.18757	0.21483
DR	500	0	0.934	0.19038	0.176863
DP	500	−14.47	22.726	0.4671	1.828188
SIZE	500	19.055	26.156	2.27E+01	1.48701
GROWTH	500	0.038	15.65	1.34697	1.647557

Table 2. Estimation results of model 1 and 2 by fixed effect and random effect models

Variables	Model	
	1	2
MO	2.612^{*}	-0.227
FO	-0.914	-0.169
IO	1.053^{*}	$.153^{*}$
BO	$.617^{*}$	-0.047^{***}
D	$.166^{**}$	-0.058^{*}
DIV	-0.007	0
Size	0.071^{**}	$.021^{*}$
Q	0.160^{*}	$.033^{*}$
R	0.481	0.083
Model type	Fix effect model	Random effect model

Estimation by FE, RE and Hausman test. Dependent variables are ACE and ACD in models 1&2 respectively. *, **, *** show significance at 1%, 5% and 10% respectively.

The contribution of manufacturing sector of Pakistan in GDP growth is slowly build up during the last three years from 18 percent to 19.2 percent. Data sources of Karachi Stock Exchange, State Bank of Pakistan, Bulletins of Statistical Bureau of Pakistan were used to collect the data (Table 1).

We checked our results by generalized least square method and Hausman Test. In models 1&2 relationship between agency cost and variables of our interest is explored. The purpose of this model is to find out the joint effect of all the variables (Table 2).

4 Results and Discussions

Managerial ownership was found influential in lowering agency cost of equity. Our results are supported by various researches [1,2,6,19,31,32,36,38,39,44,46]. Results tell that MNG is negative and significantly related with ACD. This means when managers are also owners of the firm, then they don't overlook debt-holders interests. Our results are convincing that in Pakistani firms, insiders don't create asset substitution or risk shifting problems. They don't add to the assets of owners on the expense of debt-holders. And they don't produce incentives to expropriate debt-holders' wealth as they maintain lower proportion of such assets which is not tied in fix plant and equipment or liquid assets. The motive behind this deed from managers might be because of the reason that when insiders own the firm, they want long-term performance and return from the firm, so they avoid increasing the price of debt.

Relationship between ownership of family and asset turnover of that firm is negative and significant. This means family has weak monitoring over managers. Our findings are also consistent with [37] which narrates that in family owned business, managers work for the welfare of controlling family, rather than for general shareholders. This is because Pakistan is among those emerging economies where minority rights are not well protected under any corporate law authority and due to these reason insiders with more concentration of shareholdings in companies have the capacity to misuse their rights for their personal benefits. Results are indicating that FML is negatively and significantly related with ACD i.e. increase in the ownership of family causes reduction in debt agency cost. Family owned firms have lesser proportion of such assets which are difficult to monitor by debt-holders [4].

Evidence supported that with the increase in ownership of institutions in a firm, managers make efficient utilization of assets and agency cost of equity is reduced. It can be inferred that institutions usually have strong monitoring power. Due to their monitoring and controlling power, managers are forced to make efficient utilization of assets and by this equity agency cost is reduced. Our results are consistent with majority of researchers like [1,5,6,13,46,49]. Regression results illustrate that INST is positive and significantly related with ACD. A firm which is more owned by institutions, have larger the proportion of such assets which are not tied in fixed plant and equipment, so debt-holders cannot monitor such assets easily.

With the increase in the ownership of block-holding, equity agency cost reduces. Block-holders have incentives to actively monitor agents because they have larger stake in the firm. Results tell that Pakistani block-holders have influential monitoring power. When firm is greater subjected to the monitoring of block-holders, then managers make efficient utilization of assets. So it mitigates the equity agency cost [6,42]. BLOCK is negative and significant related with debt agency cost. Higher cost of debt leads to additional cost of financing. Higher price of debt may result in the financial distress of the firm. As block-holders have large stake associated with the firm, they want to avoid such situations. So they try that there should be no debt agency cost [42,44].

Debt ratio (DR) was found to be effective in reducing both ACE and ACD. But Dividend pay out ratio came out to be insignificant in determining any type of agency cost. Size reduces ACE but on the other hand it increases the intensity of ACD. It is due to the reason that larger firms are difficult to be monitored by debt-holders. Assets of larger firms are diversified. Larger firms usually have enough free cash flows to cater future growth opportunities [13,17]. Firms having more growth, have better utilization of its assets. GROWTH is positive and highly significantly related with debt agency cost. A firm with good growth and investment opportunities is likely to have a relatively high quantity of cash and other short-term assets for future investment opportunities. The availability of more liquid assets increase the discretionary power of the manager hence increase the agency cost of debt (Table 3).

5 Conclusion

This research examines and provides further evidence on agency theory. Present study looks into the impact of ownership structures over equity agency cost (ACE) and debt agency cost (ACD). Equity agency cost is the disagreement of interest between shareholders and mangers whereas debt agency cost is the divergence of interest between shareholders and debt-holders or debt-holders and managers. Equity agency cost incurs when managers don't act in the best favor of shareholders and acquire perquisites for themselves. Agency cost of debt arises when managers, themselves or on behalf of shareholders, expropriate wealth from debt-holders and increase wealth of shareholders or do discretionary actions on the expense of debt-holders interest. Study incorporates four types of structures of ownership namely: managerial, general public, family, institutional and ownership of block-holders. We used different proxies for ACE and ACD. ACE is measured by asset turnover ratio. ACD is measured by amount of firm assets not invested in fixed plant and equipment and by liquidity of firm assets. Our sample is companies from KSE 100 index and study window is 2010–2015. We found no empirical study that has explored agency cost with set of hypotheses which present study has used.

Our research tested the convergence of interest hypothesis. In Pakistan, convergence of interest hypothesis is followed because ownership of managers' result in the reduction of agency cost. Ownership of family is causing reduction in efficient utilization of assets. However firms having managerial ownership and family ownership exhibit lower agency cost of debt. This conveys that insiders in Pakistan protect the interest of debt-holders. Due to greater monitoring abilities, institutions and block-holders have positive impact on firm performance by reducing agency cost of equity. In firms, which are more owned by institutions and block-holders, there is better protection of shareholders' interests in case of Pakistan. But institutions only take care of shareholders wealth and don't safeguard the interest of debt-holders; agency cost of debt is present in firms, which are more owned by institutions. But ownership of block-holders, not only reduces equity agency cost, but also mitigates debt agency cost by protecting debt-holders wealth and interests. So ownership of block-holders and managers are the structures of equity ownership, which reduce both types of agency costs. Pakistani firms which has greater debt ratio, also has efficient asset turnover and lower proportion of such assets which are difficult to monitor by debt-holders. Size worsens the both types of agency problems whereas growth opportunities for firm reduce agency cost of equity but exacerbates agency cost of debt.

Present study classified block-holders as all the shareholders that have more than 10% ownership of equity. Future researchers may expand this concept by classification of block-holders. For example, there are two major types of block-holders in Pakistan, family block-holders (more than 10% equity is owned by certain family) and institutional block-holders (more than 10% equity is owned by institutions). There are many types of institutions in Pakistan which have ownership in companies like NIT, pension fund managers, insurance companies etc. Future research should check individual monitoring impact of each type of institution in mitigating agency costs. The study employs five year data (2010–2015) from Pakistani economy. This period is recognized as unchecked inflation but stable or declining interest rate. Present study did not control variables for industry classification. No doubt asset utilization is different for different types of industries. Similarly other variables are different too for different industries. We were unfortunately enough that Pakistan has not codes for specific industries like SIC codes in other countries. If we use codes for different industries for control for industry specification, results might be better.

Acknowledgements. Authors of this study thank Professor Ahmad Kaleem, Dean Investment Banking Institute Business School, Melbourne for valuable suggestions.

Appendix

Table 3. Variables and definitions

Variables	Formula
Asset Turnover Ratio	Net Income divided by Total Assets
Agency Cost of Debt	Proportion of firm assets not invested in fixed plant & equipment = One minus ratio of fixed assets to total assets
Managerial Ownership	The number of common shares own by insiders divided by total number of common shares outstanding. Insiders consist of officers, affiliated directors, beneficial owners and principal shareholders [26]
Institutional Ownership	Number of shares held by institutions as a proportion from total ordinary shares of the company [36]
Family Ownership	Ratio of shares held by family as a group to total shares [43]
Ownership of Block-holders	Total number of shares held by block-holders divided by total number of shares outstanding. Where block-holders are shareholders having more than 10% ownership in total equity outstanding
Size	Natural Logarithm of total assets
Debt Ratio	Book value of contractual long term debt/Book value of total assets
Dividend	Dividend per share divided by earnings per share
Growth	Ratio of (market value of equity + book value of debt) to Book value of total assets

References

1. Agrawal A, Knoeber CR (1996) Firm performance and mechanisms to control agency problems between managers and shareholders. J Financ Quant Anal 31(3):377–397
2. Ahmed HJA (2008) Managerial ownership concentration and agency conflict using logistic regression approach: evidence from Bursa Malaysia. J Manage Res 1(1):1–10
3. Anderson RC, Reeb DM (2003) Founding-family ownership, corporate diversification, and firm leverage. J Law Econ 46(2):653–684
4. Anderson RC, Mansi SA, Reeb DM (2003) Founding family ownership and the agency cost of debt. J Financ Econ 68(2):263–285
5. Ang JS, Cox DR (1997) Controlling the agency cost of insider trading. J Financ Strateg Decis 10(1):15–26
6. Ang JS, Cole RA, Lin JW (2000) Agency costs and ownership structure. J Finance 55(1):81C106
7. Borokhovich KA, Brunarski KR et al (2005) Dividends, corporate monitors and agency costs. Financ Rev 40(1):37–65
8. Brockman P, Unlu E (2009) Dividend policy, creditor rights, and the agency costs of debt. J Financ Econ 92(2):276–299
9. Byrd JW (2010) Financial policies and the agency costs of free cash flow: evidence from the oil industry. SSRN Electron J https://ssrn.com/abstract=1664654
10. Casson M (1999) The economics of the family firm. Scand Econ Hist Rev 47(1):10–23
11. Demsetz H, Villalonga B (2001) Ownership structure and corporate performance. J Corp Finance 7(3):209–233
12. Dharwadkar B, George G, Brandes P (2000) Privatization in emerging economies: an agency theory perspective. Acad Manag Rev 25(3):650–669
13. Doukas JA, Pantzalis C (2000) Security analysis, agency costs, and company characteristics. Int Rev Financ Anal 14(5):493–507
14. Ducassy I, Guyot A (2017) Complex ownership structures, corporate governance and firm performance: the french context. Res Int Bus Finance 39:291–306
15. Faccio M, Lang LHP, Young L (2000) Dividends and expropriation. Am Econ Rev 91(1):54–78
16. Fan JPH, Wong TJ (2001) Corporate ownership structure and the informativeness of accounting earnings in East Asia. CEI Working Pap 33(3):401–425
17. Ferreira MA, Vilela AS (2004) Why do firms hold cash? evidence from EMU countries. Eur Financ Manage 10(2):295–319
18. Fleming G, Heaney R, Mccosker R (2005) Agency costs and ownership structure in Australia. Soc Sci Electron Publ 13(1):29–52
19. Florackis C (2008) Agency costs and corporate governance mechanisms: evidence for UK firms. Int J Manag Finance 4(1):37–59
20. Guariglia A, Yang J (2015) A balancing act: managing financial constraints and agency costs to minimize investment inefficiency in the Chinese market. J Corp Finance 36:111–130
21. Jensen MC (1986) Agency costs of free cash flow, corporate finance, and takeovers. Am Econ Rev 76(2):323–29
22. Jensen MC, Meckling WH (2000) The theory of the firm: managerial behavior, agency costs and ownership structure. Theory Firm, Bd 1:248–306

23. Jong AD, Dijk RV (2007) Determinants of leverage and agency problems: a regression approach with survey data. Eur J Finance 13(6):565–593
24. Jung K, Kwon SY (2002) Ownership structure and earnings informativeness: evidence from Korea. Int J Account 37(3):301–325
25. Kalcheva I, Lins KV (2007) International evidence on cash holdings and expected managerial agency problems. Rev Financ Stud 20(20):1087–1112
26. Kayakachoian G (2000) On agency costs and firms' decisions
27. Kim KA, Kitsabunnarat P, Nofsinger JR (2004) Ownership and operating performance in an emerging market: evidence from thai IPO firms. J Corp Finance 10(3):355–381
28. Kim S, Lee H, Kim J (2016) Divergent effects of external financing on technology innovation activity: Korean evidence. Technol Forecast Soc Change 106:22–30
29. Kumar J (2003) Ownership structure and corporate firm performance. Finance https://core.ac.uk/download/pdf/9315361.pdf
30. Kusnadi Y, Kusnadi Y (2004) Corporate cash holdings and corporate governance mechanisms. SSRN Electron J https://ssrn.com/abstract=479401
31. Lafond R, Roychowdhury S (2008) Managerial ownership and accounting conservatism. J Account Res 46(1):101–135
32. Lasfer M (2002) Board structure and agency costs. SSRN Electron J https://ssrn.com/abstract=314619
33. Lie E (2000) Excess funds and agency problems: an empirical study of incremental cash disbursements. Rev Financ Stud 13(1):219–247
34. Manso G (2008) Investment reversibility and agency cost of debt. Econometrica 76(2):437–442
35. Mao CX (2003) Interaction of debt agency problems and optimal capital structure: theory and evidence. J Financ Quant Anal 38(2):399–423
36. Mcknight PJ, Weir C (2009) Agency costs, corporate governance mechanisms and ownership structure in large UK publicly quoted companies: a panel data analysis. Q Rev Econ Finance 49(2):139–158
37. Morck R, Shleifer A, Vishny RW (1988) Management ownership and market valuation: an empirical analysis. J Financ Econ 20(88):293–315
38. Moussa FB, Chichti J (2011) Interactions between free cash flow, debt policy and structure of governance: 3SLS simultaneous model. J Manage Res 3(2):1–34
39. Mustapha M, Ahmad AC (2011) Agency theory and managerial ownership: evidence from Malaysia. Manag Audit J 26(5):419–436
40. Prowse SD (1990) Institutional investment patterns and corporate financial behavior in the United States and Japan. Brain Res 61(1):267–78
41. Randøy T, Goel S (2003) Ownership structure, founder leadership, and performance in norwegian SMEs: implications for financing entrepreneurial opportunities. J Bus Ventur 18(5):619–637
42. Schulze WS, Lubatkin MH, Dino RN, Buchholtz AK (2001) Agency relationship in family firms: theory and evidence. Organ Sci 12(2):99–116
43. Setia-Atmaja L, Tanewski GA, Skully M (2007) How do family ownership and control affect board structure, dividends and debt? australian evidence. In: Proceeding of The 16th European Financial Management Association Conference. Citeseer, Vienna, Austria, pp 27–30
44. Singh M, Iii WND (2003) Agency costs, ownership structure and corporate governance mechanisms. J Bank Finance 27(5):793–816
45. Stepanov S, Suvorov A (2009) Agency problem and ownership structure: Outside blockholder as a signal. SSRN Electron J 133:87–107

46. Ugurlu M (2000) Agency costs and corporate control devices in the Turkish manufacturing industry. J Econ Stud 27(6):566–599
47. Vilasuso J, Minkler A (2001) Agency costs, asset specificity, and the capital structure of the firm. J Econ Behav Organ 44(1):55–69
48. Woidtke T (2002) Agents watching agents?: evidence from pension fund ownership and firm value. J Financ Econ 63(1):99–131
49. Wu L (2004) The impact of ownership structure on debt financing of Japanese firms with the agency cost of free cash flow. SSRN Electron J
50. Yeo GHH, Tan PMS, Ho KW, Chen S (2002) Corporate ownership structure and the informativeness of earnings. J Bus Finance Account 29(7–8):1023–1046

The Impact of Institutional Investors on Stock Price Synchronicity: Evidence from the Shanghai Stock Market

Li Kun, Lei Yu(✉), and Xiaoxue Hu

Business School, Sichuan University, Chengdu 610065, People's Republic of China
245025176@qq.com

Abstract. The article mainly consists of four sections, the first section is the basic introduction and literature review, the second is the background review of institutional investors and stock synchronization, the third is the empirical research and the last is the Conclusion. This paper use the basic literature review method, descriptive statistics and OLS models analysis to explore the Impact of institutional Investors on Stock Price Synchronicity. Using the data of 569 listed firms at Shanghai stock market from 2009 to 2012, this paper investigates the impact of institutional investors' behaviors on stock price synchronicity. We find that the institutional investors' long-term investments can reduce stock price synchronicity and increase the information efficiency of the stock market. However, the short-term trading of institutional investors does not necessarily reduce stock price synchronicity. Therefore, the Chinese policemakers should regulate institutional investors and encourage constructive long-term investment behaviors, to make it really rise to enhance the effectiveness of capital market information, and improve the efficiency of capital market resource allocation. The main contributions of this paper is that we investigate the stock price synchronicity from the perspective osf institutional investors' behaviors and A–share market.

Keywords: Stock price synchronicity · Institutional investors · Long-term investment · Investment activity

1 Introduction

The term "stock price synchronicity" refers to a phenomenon in which stock prices rise and fall in tandem. As a proxy for stock price informativeness, a higher degree of synchronization reflects more commonality regarding information on markets and industry; however, heightened synchronization suggests less detailed firm-specific information contained in stock prices, weakening the capacity to transfer signals of corporate value. In this circumstance, the pricing mechanism fails, thus reducing the efficiency of resource allocation in capital markets through prices.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_130

Morck and Yeung [12] examined the stock price synchronicity in the major stock markets around the world during the first 26 weeks of 1995. They found that the average stock price synchronicity was higher in less developed countries. In the Chinese stock markets, approximately 79% of stocks move in the same direction, only lower than Poland where all the stocks moved in the same direction. After a lapse of nearly two decades, what can be observed regarding stock price synchronicity in the Chinese stock markets? Figure 1 illustrates the price synchronicity of 1,218 stocks on the Shanghai Stock Market during the first 26 weeks of 2013. As Fig. 1 shows, averagely more than 70% of stocks were moving in the same direction, suggesting that price synchronicity still remained at a high level in China.

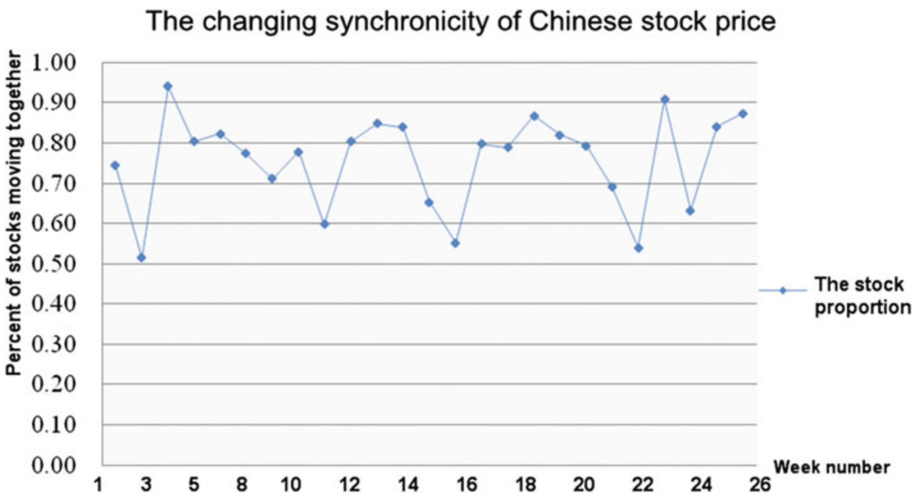


Fig. 1. The changing synchronicity of Chinese stock prices

Stock price movement is driven by investors, especially those who dominate the market. The Chinese stock market is dominated by individual investors with short-term vision. They tend to blindly follow trends without investigating the basic information of firms, thus exacerbating price synchronicity. In contrast, institutional investors are usually more professional and rational, and they engage in deeper and more fundamental research on firm-specific information. Just as Hasan and Song [6] argued, stock price synchronicity is lower when institutional investors dominate the market. However, Boubaker et al. [2] disagreed with this conclusion, they believe that the relation between institutional behavior and stock price synchronicity is not that straightforward. In recent years, institutional investors in China have been encouraged to expand for three main reasons: to return the markets to reasonable levels; to improve the pricing mechanism; and to heighten the effectiveness of market information. Therefore, it is important to clarify the influence of institutional investors on stock price synchronicity.

Since Roll, Richard [4] proposed as a proxy for stock price synchronicity, a growing body of literature has addressed this issue. Morck [12] found that the systematic component of returns variation is large in emerging markets, and lower firm-specific returns variation is associated with poorer corporate governance. Jin and Myers [3] noted that lack of transparency increases by shifting firm-specific risk to managers, and opaque stocks with high are also more likely to crash. Heng [1] found that both stock price synchronicity and crash risk are negatively related to the firm's ownership by dedicated institutional investors, which have strong incentive to monitor due to their large stake holdings and long investment horizons. However, Skaife et al. [9] argued that the zero-return metric is a better measure of firm-specific information impounded into share prices than the synchronicity measure internationally. In this paper, we use as the proxy for stock price synchronicity, because of its broad recognition and extensive applicability.

The stock price synchronicity in the Chinese stock market is significantly higher than in many other markets. Some researchers believe that this is related to the institutional development in China, e.g. Iftekhar [6] found that better institutions–respect for property rights and the rule of law - increase the value of private information which reduced synchronicity in the Chinese stock markets. However, most domestic researches tend to conclude that the high degree of speculation in the Chinese stock markets lead to high price synchronicity. This speculation is mainly caused by the composition of Chinese investors, which consists of many individuals but few institutions. This reinforces the view that institutional investors can reduce the price synchronicity. The institutional holdings can increase the quantity of firm-specific information, thus reducing stock price synchronicity and enhancing the efficiency of company signal-transmission mechanisms. However, some researchers disagree with this conclusion. Conversely, they claim that the institutional investors display "herd behavior" that can exacerbate stock price synchronicity [7].

In summary, scholars have not reached a consensus regarding the impact of institutional investors on stock price synchronicity, and more researches are required to confirm related findings, especially in the Chinese stock markets where stock price synchronicity is high. The contribution of this paper is reflected in two aspects. Firstly, this paper adds an index to demonstrate the investment activities of the institutional investors under examination. We refine institutional investor behaviors into shareholding level and examine both the effects of their long-term and short-term behaviors. This deepens the analysis of institutional investors' effect on stock price synchronicity. Secondly, this paper focuses on the study of Shanghai Stock Market, taking 569 listed companies from 2009 to 2012 as samples. The recent panel data illuminates the latest situation of stock price synchronicity in China, thus enriching the research on the effects of institutional investors on stock price synchronicity.

The remainder of this paper is organized as follows. Section 2 provides the theoretical background. Section 3 presents the research design and the main empirical results. Section 4 concludes the paper.

2 The Theoretical Background

2.1 Stock Price Synchronicity

In capital markets, the fluctuation of stock prices is driven by information shocks (unexpected or unpredictable events that affect perception of stock values). According to Roll [4], this information can be categorized into market-level and firm-level sources. The information at the market level refers to the external macro-level environment, including policies, regulations and important economic events. Firm-level information refers to the unique information of firms, including their changing financial indicators, executive turnover and other events. If the fluctuation of stock prices is caused by market-level information, then all stocks are vulnerable to the same systematic risk, and thus a significant correlation among different stock prices indicates a high degree of stock price synchronicity. As Song [10] finds, firms' stocks are less synchronized with the entire market and have less crash risk if firms have superior accounting disclosure policies. However, if the price fluctuation is based on a firms' specific information, then the entire market displays heterogeneity in terms of its stock price fluctuation. Therefore, significant stock price synchronicity in a specific market is due more to market information and due less to the firm-specific information in stock prices. On this occasion, a company's stock price cannot reflect the true quality of a company, which lowers the reference value of corporate information for investors. This damages the company signal transmission mechanism and lowers the efficiency of price-based market resource allocation.

2.2 Institutional Investors' Effects on Stock Price Synchronicity

Based on transaction incentives, traders can be classified into three categories: informed traders, noise traders, and liquidity traders. Informed traders conduct transactions based on a firm's value, and thus the related information can be conveyed into stock prices by their transactions.

Institutional investors are generally regarded as informed investors. They have access to information resources that individual investors do not have, and their teams include professional research staff. These advantages can make their evaluations of firms more accurate than those of individual investors, increasing the quantity of firm-level information contained in stock prices. Related studies have shown that the behaviors of institutional investors can affect the process of stock pricing and their information environment, and their transactions add firm-specific information into stock prices. This can help to reduce stock price synchronicity and improve the efficiency of capital markets.

The influence of institutional investor behaviors on stock price synchronicity also depends on their specific investment behaviors. On the one hand, if the institutional investors focus on short-term benefits and have high portfolio turnover rates, they will not have a strong incentive to collect firm-specific information related to corporate long-term benefits. In that case, they will not be able to decrease stock price synchronicity. On the contrary, the short-termism of institutional investors will make them more prone to herd behaviors, make them follow the trends to buy or sell stock and intensify stock price synchronicity. On the other hand, if institutional investors maintain a large stake in the firm for a long period, they may have a greater incentive to closely monitor their portfolios. To secure their investment, they may even take large positions with their shares and involve themselves in corporate governance. Therefore, the long-term investment of institutional investors is logically more likely to decrease the stock price synchronicity.

3 Empirical Design and Results

3.1 The Sample and Data

The sample consists of listed firms on the Shanghai Stock Market from 2009 to 2012, and we exclude the following firms from the sample: (1) firms without complete data; (2) financial institutions; (3) utility firms; (4) firms that conducted IPO or major assets restructuring in the current year; (5) and firms whose stocks were delisted or suspended. Finally we obtain 569 eligible companies, a total of 2276 sample points of panel data. All data are from China Stock Market & Accounting Research Database (CSMAR) and RESSET database. Table 2 shows the descriptive statistics.

3.2 Variables Construction

(1) The dependent variable: stock price synchronicity

To measure the synchronicity of stock price, we first calculate the for firm i in fiscal year t from the expanded market model regression in Eq. (1).

$$r_{i,w} = \alpha + \beta_i \times r_{m,w} + \varepsilon_{i,w}, \tag{1}$$

where $r_{i,w}$ is the return of firm i in week w, $r_{m,w}$ is the market return in week w.

Since R^2 is highly skewed and bounded between 0 and 1, we apply a logistic transformation to obtain a near formally distributed variable, SYNCH. A higher value of SYNCH indicates that the stock price is more synchronized.

$$\mathrm{SYNCH}_{i,t} = \ln\left(\frac{R^2_{i,t}}{1 - R^2_{i,t}}\right). \tag{2}$$

(2) The explanatory variables: institutional investor behavior

(a) The measurement of institutional investors' overall behaviors

We add up all the proportions of shares held by institutional investors in a listed firm to construct a proxy for the total stake of institutional investors.

$$LI_{i,t} = \frac{\sum_{j=1}^{n} \mathrm{INS}_{j,i,t}}{\mathrm{TOTAL}_{i,t}}. \tag{3}$$

The is the institutional investors' total proportion of firm i in the year t, $\mathrm{INS}_{j,i,t}$ is the shares of firm i held by institutional investor j in the year t, and $\mathrm{TOTAL}_{i,t}$ is the total shares of firm i in the year t.

(b) The measurement of institutional investors' investment liveness

According to Liu [8], 90% QFII (the Qualified Foreign Institutional Investors) hold the stocks for 1–4 quarters, and the QFII whose holding period are more than 5 quarters accounts for less than 10%. Therefore, we can deduce that the institutional investors are not necessarily long-term investors. In order to measure the short-termism of institutional investors, we construct an indicator.

$$SI_{i,t} = \frac{\mathrm{ABS}(IA_{i,t} - IA_{i,t-1})}{\mathrm{VOL}_{i,t}}, \tag{4}$$

where $IA_{i,t}$ is the total shares of firm i held by institutional investors in the year t, and $\mathrm{VOL}_{i,t}$ is the individual shareholders' trading volume of firm i in the year t. So $SI_{i,t}$ is the ratio of the absolute value of institutional investors' shares variation and stocks volume for firm i in the year t. The higher $SI_{i,t}$ means more active investment behavior of institutional investors, and short-termism is more significant.

(3) The control variables

In order to better examine the influence of institutional investors on stock price synchronicity in China, our control variables include the following: ROE, BM, SIZE, LEV, and CEN.

ROE is the contemporaneous income before extraordinary items divided by the book value of equity. ROE is a core indicator of corporate management performance, and when related to stock price synchronicity, ROE differs across models. Thus, we presently avoid making calculations on the regression coefficient.

BM is the ratio of the market value of equity to the book value of equity, and it is often used to measure the growth of a company. A lower BM ratio means stronger growth for a company. Many scholars agree that a company's growth and expanded demand for financing leads to more information disclosure; consequently more information will be contained in its stock price. In contrast, a blue-chip company with less demand for financing is reluctant to release information. Therefore, we expect that the company with a lower BM ratio will exhibit a less synchronous stock price.

SIZE is the natural log of the firms' market value of equity at the end of the fiscal year. Roll [3] held that a larger company is more vulnerable to the influence of macroeconomic or industrial events. Meanwhile, a larger company has a stronger ability to avoid risk, leading to less stock price volatility. Thus it can be concluded that the larger size of a company would bring about more significant stock price synchronicity. Others, however, hold that due to its greater scale, the company can garner more attention, which can add more information into its stock price and attenuate price synchronicity. Because of this controversy, we do not calculate a regression coefficient of the relation between SIZE and stock price synchronicity.

LEV is the book value of all liabilities scaled by total assets at the end of the fiscal year. LEV can reflect the level of a company's financial risk, and a higher LEV indicates a higher financial risk in a company. In this case, the company is more sensitive to credit risk, and thus the price fluctuation is mainly affected by the firm-specific financial risk. Thus, we expect a negative correlation between LEV and stock price synchronicity.

CEN is the concentration of corporate ownership, and a higher CEN means a stronger control held by large shareholders and more rights to internal information. These large shareholders can transfer firm-specific information into stock prices through their transactions, accordingly raising the firm-level information content in stock price [5]. However, some researchers found a concave functional relationship between stock price synchronization and CEN: CEN increases at a decreasing speed with a rising CEN until the maximum point, and then it starts to fall [11]. This suggests an unclear relationship, and any clarification would require further research.

3.3 Model Setup

We estimate the relation between institutional investors and the firms' stock price synchronicity in Shanghai A-stock market by the following model:

$$
\begin{aligned}
\text{Model (1):}\quad \mathrm{SYNCH}_{i,t} &= \alpha_t + \gamma_i + \beta_1 LI_{i,t} + \beta_3 \mathrm{ROE}_{i,t} + \beta_4 BM_{i,t} + \beta_5 \mathrm{SIZE}_{i,t} \\
&\quad + \beta_6 \mathrm{LEV}_{i,t} + \beta_7 \mathrm{CEN}_{i,t} + \varepsilon_{i,t}, \\
\text{Model (2):}\quad \mathrm{SYNCH}_{i,t} &= \alpha_t \alpha_t + \gamma_i + \beta_2 SI_{i,t} + \beta_3 \mathrm{ROE}_{i,t} + \beta_4 BM_{i,t} + \beta_5 \mathrm{SIZE}_{i,t} \\
&\quad + \beta_6 \mathrm{LEV}_{i,t} + \beta_7 \mathrm{CEN}_{i,t} + \varepsilon_{i,t}, \\
\text{Model (3):}\quad \mathrm{SYNCH}_{i,t} &= \alpha_t + \gamma_i + \beta_1 LI_{i,t} + \beta_2 SI_{i,t} + \beta_3 \mathrm{ROE}_{i,t} + \beta_4 BM_{i,t} + \beta_5 \mathrm{SIZE}_{i,t} \\
&\quad + \beta_6 \mathrm{LEV}_{i,t} + \beta_7 \mathrm{CEN}_{i,t} + \varepsilon_{i,t},
\end{aligned}
$$

where the explained variable $\mathrm{SYNCH}_{i,t}$ is the measurement of stock price synchronicity of firm i in the year t, $LI_{i,t}$ is the shares proportion held by institutional investors of firm i in the year t, and $SI_{i,t}$ is the transaction activity of firm i conducted by institutional investors during the fiscal year t, as a proxy for the short-term behavior of institutional investors.

3.4 Descriptive Statistics

Using the data of 569 listed firms at Shanghai stock market from 2009 to 2012, Table 1 provides the descriptive statistics of the sample. The mean and the max of Stock Price Synchronicity is 0.363 and 0.840 respectively. After applying a logistic transformation to obtain a near formally distributed variable, SYNCH, the mean, max and min of the SYNCH is $-0.673, 1.650$ and -5.790 respectively. The mean of LI which is the proportion of shares held by institutional investors in a listed firm is 0.155. The mean of SI which is the transaction activity conducted by institutional investors is 0.025. ROE, return on equity, show a large variation, with a mean of 4.815 and standard deviation of 17.517.

Table 1. Descriptive statistics

	N	Min	Max	Mean		Std.
	Statistic	Statistic	Statistic	Statistic	SE	Statistic
R^2	2276	.000	.840	.363	.0034	.162
SYNCH	2276	−5.790	1.650	−.673	.0188	.896
LI	2276	.000	.920	.155	.004	.169
SI	2276	.000	2.020	.025	.001	.0667
ROE	2276	−289.460	84.830	4.815	.367	17.517
BM	2276	.010	2.060	.429	.006	.280
SIZE	2276	−4.850	28.410	4.749	.215	10.262
LEV	2276	.030	.970	.539	.004	.186
CEN	2276	.000	.980	.491	.003	.165
N	2276					

3.5 Empirical Analysis

In order to decide whether we should use fixed effect panel data model or random effects model on the panel data, we conduct a Hausman Test on the foregoing model.

Table 2. Hausman test result of investment behaviour

Test Summary	Chi-Sq. Statistic	Chi-Sq.d.f	Prob.
Cross-section random (long term)	38.597	6	0.000
Cross-section random (short term)	49.000	6	0.000
Cross-section random (overall)	46.736	7	0.000

This table shows the Hausman Test result of the institutional investors' overall investment behavior. The significant result encourages us to apply a fixed-effects model in the regression process.

Table 3. The empirical result

Variable	Mod (1)	Mod (2)	Mod (3)
C	-1.135 (0.000)	-1.156 (0.000)	-1.129 (0.000)
LI	-0.487^{**} (0.037)		-0.527^{**} (0.033)
SI		-0.055 (0.846)	0.149 (0.617)
ROE	-0.005^{***} (0.001)	-0.005^{***} (0.001)	-0.005^{***} (0.001)
BM	0.871^{***} (0.000)	0.888^{***} (0.000)	0.868^{***} (0.000)
SIZE	0.004^{**} (0.013)	0.004^{**} (0.023)	0.004^{**} (0.015)
LEV	-0.502^{*} (0.070)	-0.474^{*} (0.087)	-0.498^{*} (0.072)
CEN	0.889^{*} (0.051)	0.742^{*} (0.099)	0.882^{*} (0.053)
Adjusted R-squared	0.311	0.310	0.311

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 3 reports the estimated coefficients and associated p values. From this table we find the estimated coefficient of LI is −0.487 in Model (1) and −0.527 in Model (3), and both are significant at the 5% level. This demonstrates a significantly negative correlation between the proportion of institutional shareholding and stock price synchronicity, namely, the stock with a higher proportion of holding produces a less synchronous stock price. This supports the argument that institutional investors' behavior helps to convey firm-specific information to the market and facilitate a company's signal transmission mechanism, which can improve the efficiency of resource allocation with prices.

The other explained variable, SI, is a proxy variable for institutional short-term transactions. The empirical results show a negative correlation (−0.055) in Model (2) and a positive correlation (0.149) in Model (3), neither is significant, which is consistent with expectations earlier in this paper. This indicates the insignificant correlation between the institutional investor activity and stock price synchronicity.

In this paper, we believe this insignificant correlation is caused by the double effects of institutional investors' speculative behavior. On the one hand, frequent short-term trading transmits corporate-specific information to prices, which can raise the information content in stock prices and help relieve high stock price synchronicity. On the other hand, there herding is visible in the behavior of these informed investors, causing stock price to fluctuate in the same direction. Therefore, the offsetting dual functions lead to an insignificant effect on stock price synchronicity.

Empirical result shows that ROE and BM have significant influences on stock price synchronicity, and both of them are significant at the 1% level in all the three models. Firstly, the higher ROE, the less there is price synchronicity. The firm with a higher ROE suggests better corporate management; consequently, the firm gains more attention by institutional investors and the public, and thus the firm-specific information can be readily conveyed in the price. However, the absolute value of its estimated coefficient is small, which illustrates a limited influence on price synchronicity. Secondly, the estimated coefficient of BM is significantly positive, meaning that stronger growth companies are more likely to be tracked by investors, resulting in more information in the stock price. The coefficient of SIZE is significant at the level of 5% in all three models, and this positive correlation is consistent with our former expectation. Thus a larger company suffers from more synchronous prices. The estimated coefficient of LEV and CEN are significant at the 10% level in all three models, and the stock price is less synchronous in a company with higher LEV or lower CEN.

Overall, the adjusted of the model test is approximately 31%, which indicates our model's high explanatory capability. Therefore, our conclusion is highly credible.

4 Conclusion

The empirical results of this paper suggest that institutional investment can significantly decrease stock price synchronicity on the Shanghai Stock Market, and that this influence is achieved by long-term shareholding and not by frequent transactions. This paper provides empirical grounds for encouraging institutional investors to engage in long-term value investment, which can increase the efficiency of price-based resource allocation in our capital markets. For the policy - makers, it is necessary to improve the regulations of China's capital market in order to create a favorable external environment for institutional investors. They also should guide institutional investors to long-term value investment, enhance the confidence of long-term value investment and improve the quality of listed companies.

References

1. An H, Zhang T (2013) Stock price synchronicity, crash risk, and institutional investors. J Corp Financ 21(1):1–15
2. Boubaker S, Mansali H, Rjiba H (2014) Large controlling shareholders and stock price synchronicity. J Bank Financ 40(1):80–96
3. Cfa JHG (2006) R2 around the world: new theory and new tests. J Financ Econ 79(2):257–292
4. Epps TW, Epps ML (1976) The stochastic dependence of security price changes and transaction volumes: implications for the mixture-of-distributions hypothesis. Econometrica 44(2):305–321
5. Han JF, Wang ZY (2012) Can the major shareholders trading improve market efficiency-from the perspective of the information content of stock price. J Shanxi Financ Econ Univ 7:38–45 (in Chinese)
6. Hasan I, Song L, Wachtel P (2014) Institutional development and stock price synchronicity: evidence from China. J Comp Econ 42(1):92–108
7. Jin Y, Yan M et al (2016) Stock price synchronicity and stock price crash risk: based on the mediating effect of herding behavior of qfii. China Financ Rev Int 6(3):230–244
8. Rao Y, Min L (2013) The influence of qfii holding on stock price synchronicity in China. J Manage Eng 02:202–208 (in Chinese)

9. Skaife HA, Gassen J, LaFond R (2006) Does stock price synchronicity represent firm-specific information? The international evidence. MIT Sloan Research Paper
10. Song L (2015) Accounting disclosure, stock price synchronicity and stock crash risk: an emerging-market perspective. Int J Acc Inf Manage 23(4):851–854
11. Wang X (2013) Ownership concentration and stock price synchronicity: the evidence from Chinese listed companies. Contemp Econ 14:94–95 (in Chinese)
12. Yeung B, Wu W (1999) The information content of stock markets: why do emerging markets have synchronous stock price movements? J Financ Econ 58(12):215–260

Research on Risk Allocation Model in PPP Projects: Perspectives from the SOEs

Chuan Chen, Tengteng Yan(✉), and Ping Chen

Business School, Sichuan University, Chengdu 610065, People's Republic of China
ytt19941128@foxmail.com

Abstract. PPP (Public-Private-Partnership), which is the cooperation between the government and private sector, is a major innovation in public service mechanism. Effective risk identification and reasonable risk allocation is not only a major challenge to each party, but also the key to manage projects efficiently, reduce the whole life cost and improve benefit. Under the influence of special economy, politics and culture, the state-owned enterprise accounts for absolute proportion of private sector in China. However, the systematic analysis of the state-owned enterprise participating in the PPP project is deficient. Given that the Chinese central governments have sped up the development of the PPP application in recent years, this paper provides a risk allocation model through the fuzzy comprehensive evaluation method, which is aiming at optimizing the allocation of risk, reducing transaction costs and enhancing comprehensive benefits from the perspective of state-owned enterprises (SOEs).

Keywords: PPP · SOEs · Risk allocation · Fuzzy comprehensive evaluation

1 Introduction

PPP (Public-Private-Partnership) model is an innovative financing tool for infrastructure projects procurement in China. After 2013, the 18th national congress of CPC proposed that the market should play a decisive role in resource allocation. In September 2014, the ministry of finance issued *Notice about promoting PPP Model* and devoted continuously to promoting PPP Model nationwide. In September 2015, The State Council issued *Guideline to Urge SOE Modernization through Mixed Ownership Reform*, which had further defined the overall requirements of the state-owned enterprises developing the mixed ownership economy [15]. According to the statistics of National PPP Integrated Information Platform, there were 9,285 projects that would be procured by PPP up to the end of June 20, 2016, and their total investment surged to ¥10.6 trillion [3]. Because of the national condition, the state-owned enterprises with large scale, strong capital power and rich experience have inherent advantage to participate in the PPP projects. As a result, the state-owned enterprises accounted for 80%

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_131

while the private enterprises only accounts for 20% of the private sector [14]. With a large amount of investment, long term cooperation, technical complexity, and uncertainty factors, appropriate risk allocation is not only a critical success factor (CSF) of PPP but also an important driving factor of meeting the Value for Money [4,6]. While the country issued relevant legal policies to standardize the operation and execution of PPP projects, the guidance of risk allocation is still relatively broad and vague [7].

At present, there is weak control force between government and SOEs, and government always has absolute competitive advantage. Initially, some risks would be allocated to SOEs with underperforming ability to control risk and no willingness to take risks. Once the risks occur, SOEs are unable to control the consequences of risk, thus the government need to bear losses, which inevitably causes most risks retransferred to the government at a higher cost. In addition, the complex contractual arrangements in PPP projects would lead to risk exposure, and inappropriate risk allocation only let the infrastructure debt simply transfer from the local government balance sheets to SOEs. However, it can't cut overall leverage utility of the public sector effectively. In terms of risk allocation, the government and SOEs always face severe challenge, strict requirements and high standards. However, there are few systematic analysis of the state-owned enterprises participating in the PPP project in Chinese journals. Therefore, this paper will establish risk allocation model combined with fuzzy comprehensive evaluation method and entropy coefficient method to optimize the allocation of risk perspectives from the SOEs.

This paper is organized as the following: Sect. 1 begins with an introduction, Sect. 2 makes a literature review about PPP risk allocation, Sect. 3 describes the modeling ideas and methods combined with fuzzy comprehensive evaluation method and entropy coefficient method, which optimizes multi-objective decision of risk allocation in PPP projects perspectives from the SOEs, Sect. 4 takes the case of *Mianzhu Integration of Water Supply and Drainage Project* to demonstrate the effectiveness of the model, Sect. 5 presents a brief conclusion .

2 Literature Review

Zhang put forward that the research topics of PPP papers in Chinese and International Journals mainly include PPP model's application, risk management, financing and economic issues, legal and procurement issues and government regulation [16]. The foundation of financing risk-sharing in PPP projects is the accurate identification of risk factors. Grimsey identified nine risks: financial risk, political risk, environment risk, construction risk, technology risk, operational risk of infrastructure, infrastructure recovering risks, force majeure and project default risk, introduced the risk evaluation index and evaluation method from the perspective of different project stakeholders [9]. Bing creatively carried out the dividing analysis of risk from the macro, medium and micro three levels [2]. Zou identified twenty-seven critical risk factors influencing the success of PPP projects from five dimensions: the owner, design, contractors, government institutions and the external environment [17].

The academic research for the risk-sharing mechanism in PPP project mainly concentrated on the risk-sharing principle, method and model. Abednego put forward the connotative meaning and conditions of risk sharing [1]. Deng analyzed and summarized nine principles in risk-sharing of PPP projects, such as the principle of responsibility, justice principle [8]. Bing put forward that the risk-bearer should be divided into the private sector, public sector, together undertaking, which is depended on the specific conditions of project [2]. Jin and Chang explained the risk allocation of PPP projects from the perspective of transaction cost theory (TCE) and resource-based view (RBV) [5,11]. At present, most studies tend to adopt quantitative methods to optimize the PPP project risk sharing mechanism and process. The questionnaire survey was the typical quantitative research method in risk allocation [16], such as Ke investigated the experienced practitioners to determine the preferences of PPP project risk allocation [13], Guo adopted the real option method to research the risk allocation strategy of highway PPP projects under the delay investment decisions [10], Jin selected the artificial neural network to realize the allocation [12].

In conclusion, with the perfection and development of theory and practice of risk-sharing in PPP projects, the scholars at home and abroad had a thorough study in risk allocation, which mostly confined to some specific risks that should be shared to specific stakeholders, government or private sector. However, there are few studies involved specific allocation strategies and proportions. Moreover, there is almost no study about the risk allocation whose perspective is from the SOEs.

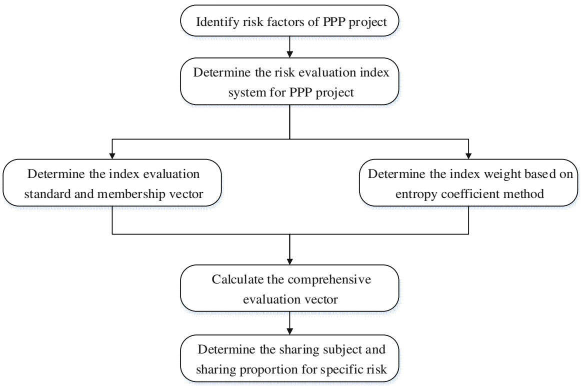


Fig. 1. The procedure of risk allocation whose perspective from the SOEs in PPP project

3 Method

The quantitative method adopted in optimizing risk allocation mainly include game theory, real option theory, artificial neural networks and fuzzy comprehensive evaluation theory. The game theory, option theory and artificial neural networks objectively reflect the actual negotiation process with complex model, which are difficult to apply in practice. In consideration of subjectivity and uncertainty in risk-decision mechanism, the fuzzy comprehensive evaluation (FCE) based on fuzzy mathematics will be reasonable in optimizing risk allocation. It can turn the qualitative questions into quantitative analysis based on membership degree theory. In addition, the method has the advantages of clear result, quite strong systematicness and can better solute the problems that are difficult to quantify, which has been widely applied in the field of engineering and applied science, such as water quality evaluation, clinical diagnosis, information technology, air quality assessment and product design. Some scholars also gradually attempt to apply fuzzy mathematics in the construction risk management. The paper establishes the PPP project risk allocation model to optimize the multi-objective decision-making in PPP project risk allocation perspectives from the SOEs. Figure 1 summarizes the procedure of PPP project risk allocation.

3.1 Identify Risk Factors of PPP Project

Risk identification further defined the risk factors and risk category, which is the foundation of risk allocation. In the practice of risk management in PPP project, there are several applicable methods, such as the *Delphi* method, Check list method and Brainstorming method. There are not many differences in risk factors and risk category between the SOEs and private enterprise, but there are considerable differences in the specific risk assessment. Moody's Investors Service said that China's state-owned enterprises were heavily involved in the PPP project, and it would reduce government leverage and debt at a certain extent. With the promoting of supply-side structural reform, Moody's said that the proportion of the debt in GDP in SOEs reaches to approximate 115%. Mostly SOEs faced systemic downgrade in credit rating, these factors would affect the risk assessment of PPP project.

3.2 Determine the Risk Evaluation Index System for PPP Project

This section discusses the risk evaluation index system based on allocation principle and influence factors. There are many factors and principles affecting the risk allocation of state-owned enterprises, the paper selects the basics of qualitative research to comb the influencing factors and principles. The influencing factors include project system properties, stakeholders' comprehending mistake for the PPP model, the risk-taking attitude and intention of stakeholders. Combined with the theoretical research and case study, the guiding principle of PPP project could be summarized as follows: equitable principle, criterion of liability, effective control principle, risk premium peer, upper limit principle and dynamic principle. Setting up the evaluation index system should be carried out in accordance with the principle: quantifiable, critical, objectivity, systematic, operability, independence, comparability and representative, which can full and accurately reflect the comprehensive management ability to control risk of the specific subject. Table 1 shows the evaluation index system.

Table 1. The evaluation index system in risk allocation

The evaluation index system			
Ability to control risk	U1	Foresee Risk	It can predict the probability and evaluate the consequences of risk occurrence (Degree of information symmetry)
	U2	Ability to monitor risk	The ability to avoid, reduce, supervise and control the likelihood of occurrence for the specific Risk
	U3	Ability to improve the risk management system	Construct management information system, improve the measures of dealing with risks
Ability to undertake risks	U4	The flexibility of dealing with risk	The ability to stabilize and control risk consequences, don't make it worse, take risks in the lowest price
	U5	Solid financial performance	Private capital (asset-liability ratio, risk reserve rate, profitability, etc.) the government (fiscal balances, GDP growth, debt levels)
	U6	The experience to cope with the risk	Rich experience and reputation of honoring agreements makes better coordinate various stakeholders, and then take risks at the lowest cost
Risk premium level	U7	Obtain direct premium level	The ability to get a reasonable and acceptable level for risk premium
	U8	Obtain indirect premium level	The ability to improve the credibility, reputation and efficiency of stakeholders
Risk preference	U9	Risk attitude	The tendency of the subject facing the specific risks, such as risk-averse, risk-neutral, risk-taking
	U10	Inclination of transfer risk	Transfer inclination based on the ownership and control power between the SOEs and the government

3.3 Determine the Index Evaluation Standard and Membership Vector

The fuzzy comprehensive evaluation model is based on the franchise agreement between the government and private sector. Combined with the risk evaluation index system in Table 1 and triangular fuzzy set theory, this paper set up the fuzzy reasoning rules. V includes five evaluation standards, $V =$ (Very low, Low, Medium, High, Very high).

The triangular fuzzy function $T = (f, g, h)$, Very low (0%, 15%, 30%), Low (15%, 25%, 35%), Medium (35%, 45%, 55%), High (55%, 65%, 75%), Very high (75%, 85%, 95%), Through the formula $V(T) = \frac{1}{3}(f + g + h)$ to calculate the fuzzy evaluation of value $V =$ (15%, 25%, 45%, 65%, 85%).

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{pmatrix}.$$

We use *Delphi* method to determine the evaluation value. This paper establishes the subordinated matrix R based on the index system and evaluation standard to calculate the subordinated vector $r_i = (r_{i1}, r_{i2}, \cdots, r_{im})$ indicates the value of index i in the evaluation standard of j.

3.4 Determine the Index Weight Based on Entropy Coefficient Method

Considering the subjectivity of expert evaluation system, this paper adopts entropy coefficient method to determine the index weight, thus enhancing the reliability of the model and improving the accuracy and rationality of the evaluation results. The entropy concept derives from thermodynamics, which initially described the system disorder. Information entropy value is directly proportional to the disorder, which is inversely proportional to the information utility value. Entropy coefficient method is mainly according to the principle of $VarianceDrive$, which can largely avoid human factors to determine the index weights. The specific steps are as follows:

(1) Use the Stirling Formula to calculate the entropy value of each index utilizing Stirling Formula.

$$e_i = -\frac{1}{Inm}\sum_{j=1}^{m} R_{ij} InR_{ij}.$$

(2) Determine the weight value of each index.
- Utility value h_i is the difference between the entropy e_i and 1. $h_i = 1 - e_i$;
- The weight value of each index can be indicated as $W = (w_1, w_2, w_3 \cdots w_i)$;

$$w_i = \frac{h_i}{\sum_{i=1}^{n} h_i}.$$

3.5 Calculate the Comprehensive Evaluation Vector

According to subordinated vector and the index weight, we need to calculate the comprehensive evaluation vector $B = W \circ R$, $\circ$ represents fuzzy operator,

Table 2. Alternative fuzzy operators

Characteristic	Alternative fuzzy operators			
	$M(\wedge,\vee)$	$M(\bullet,\vee)$	$M(\wedge,\oplus)$	$M(\bullet,\oplus)$
Formula	$\max\limits_{1\le i\le n}\{\min(w_i,r_{ij})\}$	$\max\limits_{1\le i\le n}\{\min(w_i,r_{ij})\}$	$\min\left\{1,\sum\limits_{i=1}^{n}\min(w_i,r_{ij})\right\}$	$\min\left\{1,\sum\limits_{i=1}^{n}w_i r_{ij})\right\}$
Manifests the weighting function	Not obvious	obvious	Not obvious	obvious
Comprehensive degree	Weak	Weak	Strong	Strong
Using R information	Inadequate	Inadequate	More adequate	Adequate
Type	Dominant-factor Highlight	Dominant-factor Highlight	Weighted average model	Weighted average model

the commonly used fuzzy operators include $M(\wedge,\vee)$, $M(\bullet,\vee)$, $M(\wedge,\oplus)$, $M(\bullet,\oplus)$, $M(\bullet,\oplus)$ comprehensively consider the subordinated vector and weight of the index, and other fuzzy operator would lack some information $M(\wedge,\vee)$, $M(\bullet,\vee)$ are suitable for single constraints model, which only take the main indicators into account and ignore the other secondary index. The fuzzy operators' characteristics are shown in Table 2.

Based on the above analysis, this paper will choose $M(\bullet,\oplus)$to calculate the comprehensive evaluation vector $B = (b_1, b_2, b_3 \cdots b_m)$. The methods widely adopted to deal with the evaluation vector are weighted average method and the principle of maximum membership degree. However, the latter take the concept of optimal sorting in the evaluation scheme, and sometimes it would miss key information. So this section select the weighted average method to normalize the index system, which makes the comprehensive evaluation vector $B = (b_1, b_2, b_3, \cdots, b_m)$ converted to comprehensive score S. $S = B_K \times V$, V represent the weight of e index evaluation standard, $V = (0.150.250.450.650.85)$.

3.6 Determine the Sharing Subject and Sharing Proportion for Specific Risk

The government and SOEs respectively undertake a comprehensive assessment of specific risk, the government should bear the proportion $Q_1 = \frac{S_1}{S_1+S_2} \times 100\%$, SOEs should bear the proportion $Q_2 = \frac{S_2}{S_1+S_2} \times 100\%$.

4 Case Study

To further integrate the existing system of water supply and drainage, improve the systematicness, harmonization, sharing and economy of water supply and drainage, the government of Mianzhu, Sichuan province proposed to launch the Mianzhu Integration of Water Supply and Drainage project in November 2015. The project's investment is 680 million yuan, and its concession period is about 30 years, which includes the construction and reform period of 3 years. Mianzhu

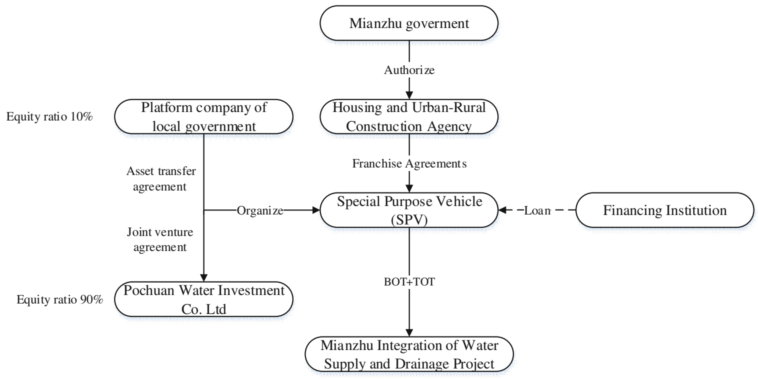


Fig. 2. The investment and financing structure

government authorized Housing and Urban-Rural Construction Agency as implementing agency to supervise and manage the implementation of project. The project intended to adopt BOT & TOT mode and complete government procurement through competitive consultation on February 5, 2016. Eventually, Pochuan Water (SOEs) was selected, and then the two sides formally signed on June 14, 2016. Pochuan Water is a joint-stock subsidiary companies of Sichuan Development Equity Investment and Fund Management Co. Ltd. and Poten Environment Group Co. Ltd. The investment and financing structure is shown in Fig. 2.

4.1 Risk Identification List Based on the VFM Evaluation

According to *Guidelines for VFM Evaluations of PPP (for Trial Implementation)* (Cai Jin [2015] No. 167), the VFM evaluation meeting was convened. The risk list identified by the participating experts is shown in Table 3.

Take the tax regulation change for example, this section will verify the application of fuzzy comprehensive evaluation method in the risk allocation decisions between Mianzhu government and Pochuan Water.

4.2 Determine the Subordinated Vectors of Mianzhu Government and Pochuan Water

According to the evaluation score from 30 experts (project management, engineering, finance, assets evaluation, industry, water industry, law, registered accounting field), the subordinated vectors of Mianzhu government and Pochuan Water are R_1, R_2. It shows that the comprehensive assessment about the ability to control and undertake risk, incentive mechanism, risk preference for Mianzhu government and Pochuan Water according to Mianzhu Integration of Water

Table 3. Risk identification list of Mianzhu Integration of Water Supply and Drainage project

Risk category	Major risks	Developing stage
Financial risk	Capital availability	Construction
	Credit risk in financial institutions	Construction
	Exchange rate risks	Construction and Operation
	Interest rate risk	Construction and Operation
	Inflation Risk	Operation
	Refinancing risk	Operation
	liquidity risk	Construction and Operation
Design risk	Design changeThe government put forward	Construction
	Design changeThe SOEs put forward	Construction
Construction risk	Geological exploration risk	Construction
	Low process/technology	Construction
	Duration delay	Construction
	Cost overrun	Construction
	Not timely access to labor/equipment	Construction
	Acquire land	Construction
	Project quality	Construction
	Construction safety	Construction
	Damaged environment and cultural relics	Construction
Operational risk	Operation quality	Operation
	Operation safety	Operation
	Payment delay	Operation
	Water quantity and water quality change risk	Operation
Law and policy	National law/policy changes	Full life-cycle
Tax regulation change	National general tax or tax law applicability	Operation
Political risks	Approval delay	Construction
Transfer risk	Not meet the requirements of handover	Transfer
Natural environmental risk	Force majeure	Full life-cycle
Residual risk	Not considering the unknown risk	Full life-cycle

Supply and Drainage project. R_1, R_2 are shown below.

$$R_1 = \begin{bmatrix} 0.50 & 0.23 & 0.20 & 0.04 & 0.03 \\ 0.15 & 0.42 & 0.24 & 0.13 & 0.06 \\ 0.01 & 0.03 & 0.26 & 0.40 & 0.31 \\ 0.59 & 0.32 & 0.03 & 0.03 & 0.03 \\ 0.01 & 0.01 & 0.22 & 0.50 & 0.26 \\ 0.08 & 0.04 & 0.53 & 0.26 & 0.09 \\ 0.02 & 0.14 & 0.37 & 0.34 & 0.14 \\ 0.02 & 0.01 & 0.53 & 0.40 & 0.04 \\ 0.06 & 0.13 & 0.52 & 0.23 & 0.06 \\ 0.33 & 0.27 & 0.14 & 0.17 & 0.09 \end{bmatrix}, \quad R_2 = \begin{bmatrix} 0.29 & 0.34 & 0.25 & 0.12 & 0.01 \\ 0.02 & 0.10 & 0.37 & 0.34 & 0.17 \\ 0.00 & 0.05 & 0.44 & 0.34 & 0.17 \\ 0.05 & 0.20 & 0.27 & 0.24 & 0.24 \\ 0.07 & 0.12 & 0.49 & 0.25 & 0.07 \\ 0.01 & 0.05 & 0.49 & 0.32 & 0.14 \\ 0.07 & 0.10 & 0.51 & 0.25 & 0.07 \\ 0.01 & 0.07 & 0.59 & 0.24 & 0.10 \\ 0.01 & 0.05 & 0.65 & 0.22 & 0.07 \\ 0.08 & 0.21 & 0.37 & 0.21 & 0.13 \end{bmatrix}.$$

4.3 Determine the Index Weight Based on Entropy Coefficient Method

According to the entropy coefficient method, the index weight is calculated following the step 4 in the above chapters.

(1) For Mianzhu government

$$InR_{ij} = \begin{bmatrix} -0.69 & -1.47 & -1.61 & -3.22 & -3.51 \\ -1.90 & -0.87 & -1.43 & -2.04 & -2.81 \\ -5.30 & -3.51 & -1.37 & -0.92 & -1.17 \\ -0.53 & -1.14 & -3.51 & -3.51 & -3.51 \\ -4.31 & -4.27 & -1.54 & -0.69 & -1.36 \\ -2.53 & -3.22 & -0.63 & -1.35 & -2.41 \\ -4.20 & -2.00 & -0.99 & -1.08 & -1.97 \\ -3.91 & -4.61 & -0.63 & -0.92 & -3.22 \\ -2.81 & -2.04 & -0.65 & -1.47 & -2.81 \\ -1.11 & -1.31 & -1.97 & -1.77 & -2.41 \end{bmatrix},$$

$$R_{ij} \bullet InR_{ij} = \begin{bmatrix} -0.35 & -0.34 & -0.32 & -0.13 & -0.11 \\ -0.28 & -0.36 & -0.34 & -0.27 & -0.17 \\ -0.03 & -0.11 & -0.35 & -0.37 & -0.36 \\ -0.31 & -0.36 & -0.11 & -0.11 & -0.11 \\ -0.06 & -0.06 & -0.33 & -0.35 & -0.35 \\ -0.20 & -0.13 & -0.34 & -0.35 & -0.22 \\ -0.06 & -0.27 & -0.37 & -0.37 & -0.28 \\ -0.08 & -0.05 & -0.34 & -0.37 & -0.13 \\ -0.17 & -0.27 & -0.34 & -0.34 & -0.17 \\ -0.37 & -0.35 & -0.28 & -0.30 & -0.22 \end{bmatrix},$$

$$e_i = -\frac{1}{Inm}\sum_{j=1}^{m} R_{ij}InR_{ij} = \begin{bmatrix} 0.77 & 0.89 & 0.75 & 0.62 & 0.71 & 0.77 & 0.83 & 0.59 & 0.80 & 0.94 \end{bmatrix},$$

$$h_i = 1 - e_i = \begin{bmatrix} 0.23 & 0.11 & 0.25 & 0.38 & 0.29 & 0.23 & 0.17 & 0.41 & 0.20 & 0.06 \end{bmatrix},$$

$$w_i = \frac{h_i}{\sum_{i=1}^{n} h_i} = \begin{bmatrix} 0.10 & 0.05 & 0.11 & 0.16 & 0.12 & 0.10 & 0.07 & 0.17 & 0.09 & 0.03 \end{bmatrix}.$$

(2) For Pochuan Water

$$InR_{ij} = \begin{bmatrix} -1.24 & -1.08 & -1.39 & -2.16 & -5.30 \\ -3.91 & -2.30 & -0.99 & -1.08 & -1.77 \\ -6.91 & -3.02 & -0.82 & -1.08 & -1.77 \\ -3.00 & -1.61 & -1.31 & -1.43 & -1.43 \\ -2.66 & -2.12 & -0.71 & -1.39 & -2.66 \\ -5.30 & -3.10 & -0.71 & -1.14 & -1.97 \\ -2.66 & -2.30 & -0.67 & -1.39 & -2.66 \\ -5.30 & -2.66 & -0.54 & -1.43 & -2.30 \\ -4.61 & -3.00 & -0.43 & -1.51 & -2.66 \\ -2.53 & -1.56 & -0.99 & -1.56 & -2.04 \end{bmatrix},$$

$$R_{ij}InR_{ij} = \begin{bmatrix} -0.36 & -0.37 & -0.35 & -0.25 & -0.03 \\ -0.08 & -0.23 & -0.37 & -0.37 & -0.30 \\ -0.01 & -0.15 & -0.36 & -0.37 & -0.30 \\ -0.15 & -0.32 & -0.35 & -0.34 & -0.34 \\ -0.19 & -0.25 & -0.35 & -0.35 & -0.19 \\ -0.03 & -0.14 & -0.35 & -0.36 & -0.28 \\ -0.19 & -0.23 & -0.34 & -0.35 & -0.19 \\ -0.03 & -0.19 & -0.31 & -0.34 & -0.23 \\ -0.05 & -0.15 & -0.28 & -0.33 & -0.19 \\ -0.20 & -0.33 & -0.37 & -0.33 & -0.27 \end{bmatrix},$$

$$e_i = -\frac{1}{Inm}\sum_{j=1}^{m} R_{ij}InR_{ij} = \left[0.84\ 0.84\ 0.74\ 0.94\ 0.82\ 0.72\ 0.80\ 0.68\ 0.62\ 0.93\right],$$

$$h_i = 1 - e_i = \left[0.16\ 0.16\ 0.26\ 0.06\ 0.18\ 0.28\ 0.20\ 0.32\ 0.38\ 0.07\right],$$

$$w_i = \frac{h_i}{\sum_{i=1}^{n} h_i} = \left[0.08\ 0.08\ 0.13\ 0.03\ 0.09\ 0.14\ 0.09\ 0.15\ 0.18\ 0.04\right].$$

Comprehensively considering the Mianzhu government and Pochuan Waters' information entropy evaluated by 30 experts, finally it is concluded that the weight of each index value are calculated. $W = [0.09\quad 0.06\quad 0.12\quad 0.10\quad 0.10\quad 0.12\quad 0.08\quad 0.16\quad 0.14\quad 0.03]$.

4.4 Calculate the Comprehensive Evaluation Vector (Comprehensive Membership)

Based on the above calculation, using fuzzy operator can calculate the comprehensive evaluation vector for the Mianzhu government and Pochuan Water, and the results are shown below.

$$
\begin{aligned}
B_1 &= W \circ R_1 \\
&= \begin{bmatrix} 0.09 & 0.06 & 0.12 & 0.10 & 0.10 & 0.12 & 0.08 & 0.16 & 0.14 & 0.03 \end{bmatrix} \circ \begin{bmatrix} 0.50 & 0.23 & 0.20 & 0.04 & 0.03 \\ 0.15 & 0.42 & 0.24 & 0.13 & 0.06 \\ 0.01 & 0.03 & 0.26 & 0.40 & 0.31 \\ 0.59 & 0.32 & 0.03 & 0.03 & 0.03 \\ 0.01 & 0.01 & 0.22 & 0.50 & 0.26 \\ 0.08 & 0.04 & 0.53 & 0.26 & 0.09 \\ 0.02 & 0.14 & 0.37 & 0.34 & 0.14 \\ 0.02 & 0.01 & 0.53 & 0.40 & 0.04 \\ 0.06 & 0.13 & 0.52 & 0.23 & 0.06 \\ 0.33 & 0.27 & 0.14 & 0.17 & 0.09 \end{bmatrix} \\
&= \begin{bmatrix} 0.15 & 0.13 & 0.34 & 0.27 & 0.11 \end{bmatrix}, \\
B_2 &= W \circ R_2 \\
&= \begin{bmatrix} 0.09 & 0.06 & 0.12 & 0.10 & 0.10 & 0.12 & 0.08 & 0.16 & 0.14 & 0.03 \end{bmatrix} \circ \begin{bmatrix} 0.29 & 0.34 & 0.25 & 0.12 & 0.01 \\ 0.02 & 0.10 & 0.37 & 0.34 & 0.17 \\ 0.00 & 0.05 & 0.44 & 0.34 & 0.17 \\ 0.05 & 0.20 & 0.27 & 0.24 & 0.24 \\ 0.07 & 0.12 & 0.49 & 0.25 & 0.07 \\ 0.01 & 0.05 & 0.49 & 0.32 & 0.14 \\ 0.07 & 0.10 & 0.51 & 0.25 & 0.07 \\ 0.01 & 0.07 & 0.59 & 0.24 & 0.10 \\ 0.01 & 0.05 & 0.65 & 0.22 & 0.07 \\ 0.08 & 0.21 & 0.37 & 0.21 & 0.13 \end{bmatrix} \\
&= \begin{bmatrix} 0.05 & 0.11 & 0.47 & 0.25 & 0.11 \end{bmatrix}.
\end{aligned}
$$

4.5 Determine the Sharing Subject and Sharing Proportion for Tax Regulation Change Risk

Then the section normalize the index system, which makes the comprehensive evaluation vector $B = (b_1, b_2, b_3, \cdots, b_m)$ converted to comprehensive score S.

$$
\begin{aligned}
S_1 &= B_1 \times V = \begin{bmatrix} 0.15 & 0.13 & 0.34 & 0.27 & 0.11 \end{bmatrix} \times \begin{bmatrix} 0.15 & 0.25 & 0.45 & 0.65 & 0.85 \end{bmatrix}^T = 0.48080, \\
S_2 &= B_2 \times V = \begin{bmatrix} 0.05 & 0.11 & 0.47 & 0.25 & 0.11 \end{bmatrix} \times \begin{bmatrix} 0.15 & 0.25 & 0.45 & 0.65 & 0.85 \end{bmatrix}^T = 0.50897.
\end{aligned}
$$

The Mianzhu government should bear the proportion:

$$Q_1 = \frac{S_1}{S_1 + S_2} \times 100\% = 48.577\%.$$

The Pochuan Water should bear the proportion:

$$Q_2 = \frac{S_2}{S_1 + S_2} \times 100\% = 51.423\%.$$

The example shows that the risk of tax regulation change should be shared by the Mianzhu government and Pochuan Water. The shared proportions were 48.577% and 51.423%. The Tax regulation change is belong to macroeconomic level risk, and the local government or superior government, central government can influence macroeconomic conditions. But the allocation is involved in local government and social capital, the local government is unable to fully control the occurrence and adverse consequences adverse consequences of tax risk, therefore the tax regulation risk should be shared.

5 Conclusion

Risk allocation is the key factor in the success of infrastructure PPP projects, and the reasonable risk allocation can effectively reduce the risk level. This paper explores the effective approach of risk allocation form the perspectives of the SOEs, and builds the allocation model based on fuzzy comprehensive evaluation and entropy coefficient method. In addition, the paper demonstrates the objectivity, rationality and feasibility of the allocation model, which provides the scientific methods and guidance for SOEs to participate in PPP projects, further deepening the reform of state-owned enterprises and gradually making the economic grows change from the factor-driven into the innovation-driven, promoting the transformation of public ownership economy.

Acknowledgements. This paper is attributed to the financial support from the National Natural Science Foundation of China (NSFC) under the project No. 71371131.

References

1. Abednego MP, Ogunlana SO (2006) Good project governance for proper risk allocation in public-private partnerships in Indonesia. Int J Project Manage 24(7):622–634
2. Bing L, Akintoye A et al (2005) The allocation of risk in PPP/PFI construction projects in the UK. Int J Project Manage 23(1):25–35
3. China Public Private Partnerships Center (2016) Quarterly report on the project library of the national PPP integrated information platform. http://www.cpppc.org/en/Quarterly/4007.jhtml
4. Chan APC, Lam PTI (2010) Critical success factors for PPPs in infrastructure developments: Chinese perspective. J Constr Eng Manage 136(5):484–494
5. Chang C (2013) A critical review of the application of TCE in the interpretation of risk allocation in PPP contracts. Constr Manage Econ 31(2):99–103
6. Chen C, Chen P, Wang Q (2017) Comparing the efficiency of public-private partnerships with the traditional procurement: based on the Chengdu No. 6 water plant B. In: Proceedings of the tenth international conference on management science and engineering management, pp 487–501
7. Chou JS, Pramudawardhani D (2015) Cross-country comparisons of key drivers, critical success factors and risk allocation for public-private partnership projects. Int J Project Manage 33(5):1136–1150
8. Deng X, Qiming LI et al (2008) Summary and application of the principles of risk allocation in PPP model. Constr Econ 09:32–35 (in Chinese)
9. Grimsey D, Lewis MK (2002) Evaluating the risks of public private partnerships for infrastructure projects. Int J Project Manage 20(2):107–118

10. Jian G (2013) The risk allocation strategy research of PPP project in highway traffic infrastructure. Manage Rev 25(7):11–19
11. Jin XH (2009) Determinants of efficient risk allocation in privately financed public infrastructure projects in Australia. J Constr Eng Manage 136(2):138–150
12. Jin XH, Zhang G (2011) Modelling optimal risk allocation in PPP projects using artificial neural networks. Int J Project Manage 29(5):591–603
13. Ke Y, Wang SQ et al (2010) Preferred risk allocation in China's public-private partnership (PPP) projects. Int J Project Manage 28(5):482–492
14. Liu S (2016) Promote the reform of state-owned enterprises in the PPP way. Mod SOE Res 9:78–81 (in Chinses)
15. The State Council (2015) China urges SOE modernization through mixed ownership reform. http://english.gov.cn/policies/latest_releases/2015/09/24/content_281475197422388.htm
16. Zhang S, Chan APC et al (2016) Critical review on PPP research-a search from the Chinese and international journals. Int J Project Manage 34(4):597–612
17. Zou PXW, Zhang G, Wang J (2007) Understanding the key risks in construction projects in China. Int J Project Manage 25(6):601–614

An Exploratory Case Study of the Mature Enterprise's Corporate Brand Building Based on Strategic Perspective

Yuan Yuan, Jing Xu, Liming Zhang(✉), and Rui Zhou

Business School, Sichuan University, Chengdu 610064, People's Republic of China
zhangliming@scu.edu.cn

Abstract. Corporate brand plays an increasingly important role in intensive competition. Therefore, further research needs to be done on the relationship between strategic action and corporate brand. Particularly for mature enterprises, how to build corporate brand through strategic action is an important issue in this research field. Based on the case studies of Jinjiang Hotel Industrial Co., Ltd. and Luzhou Laojiao Co., Ltd., this paper analyzes the corporate brand management of mature manufacturing and service-oriented enterprises, giving the conclusion that strategic action will contribute to the corporate brand management and the brand construction of a mature enterprise is the consequence of its strategic action, since enterprises can manage corporate brand building by improving brand awareness, brand structure and the building path to construct their own corporate brands.

Keywords: Corporate brand · Strategic action · Mature enterprise

1 Introduction

Over the past two decades, corporate branding has been advocated by scholars as an effective way for companies to build competitive advantage [4], it has become a key strategic resource of enterprise [9]. In the contemporary global competitive environment, products and services typically become more similar, which makes it difficult for consumers to differentiate offerings coming from different enterprises. Consequently, promoting entire companies as a brand has become an efficient approach to create differentiation [8]. Corporate brand has significant value in consumers concern, product supporting, core communication, employee incentive [13]. And it is capable of helping enterprise to reduce cost, giving customers a sense of belonging, building consensus in stakeholders [11]. Therefore, the study on corporate brand from the perspective of strategy is significant.

To our knowledge, existing research about corporate brand mainly focuses on corporate brand itself [3,5,10,12,14], or on the marketing area, which explores

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_132

how corporate brand affects the consumers' reviews, attitudes or purchasing tendency [3,5,15,16]. Besides, the relationships among corporate brand and organization identity, image, reputation, etc. [7,12], the influence factors of corporate brand, and the effect of corporate brand on other aspects [7,8,16,17] are also research hotspots now. But there is little research on the inner relationship between corporate brand and enterprise strategy. Especially for mature enterprises, as to promote the corporate brand, the ways of facing the competition from start-ups and releasing from past traces become essential issues to be resolved.

So, this research plans to carry on an exploratory case study of two mature company cases, to analyze the role and impact of strategic actions taken in the enterprise development process on corporate brand rebuilding.

2 Literature Review

The research on corporate brand sprang up in 1930s'. To some extent, corporate brand is defined as an organization from the whole, which is the platform of all corporate brands and carrier used for promoting various differential services [1]. However, Knox and Bickerton [12] thought corporate brand is a manifestation of unique business mode of an organization in vision, diffusion and behavior. Harch and Schultz [11] defined corporate brand as brand umbrella covering all corporate products and plate brilliance for product.

2.1 Elements of Corporate Brand

Aaker [1] concluded the elements of corporate brand are historical foundation, corporate performance, employee performance, value view and priority item, local and global orientation, social resonance. Hatch and Schultz [9] proposed that corporate brand is made up of vision, culture and image. Vision comes from leadership consciousness of senior leaders, culture comes from employees, forming internal cognition and cohesion, and image is the evaluation of external public. Balmer [6,10] pointed out the components of corporate brand include brand vision, culture, location, personal feature, public relation and information.

2.2 Brand Building and Maintenance

Brand building and maintenance work need to design systematic and clear guidance plan from organizational strategic level. Philip Kolter [2] thought that enterprise should recognize advantages and disadvantages of brand regularly, maintain the development of brand constantly and propose four aspects of brand construction which are positioning, name selection, holder's decision and development strategy. With the gradual increasing of competition, corporation pays more attention to seek for differential competition of brand through unique emotional experience rather than function or characteristic in brand construction [10].

2.3 Management and Application of Corporate Brand

It's clear that theoretical research on definition, composition, role and feature, etc. of corporate brand at present has made some achievement, but corporate brand management research is somewhat lacking in guiding practice. Mukherjee and Balmer [14] clearly pointed out that the current articles about corporate brand are too ideal, they lack empirical data support and related measure description for possible conditions. The news shows that many scholars are developing related research on corporate branding, for example, Knox and Bickerton [12] pointed out corporate brand management is suggested to consider four aspects as vision management, culture management, image management and competition management. Hatch and Schultz [11] proposed corporate brand management is the process of eliminating the difference of corporate vision, culture and image.

In conclusion, not only does corporate brand has profound relationship with corporate strategic action, but it has close relation with corporate development stage and business type. Strategic action is strategic operation affecting future development direction of corporation and the result of executing corporate strategic plan, which reflects corporate strategic choice. However, strategic choice will promote the formation and change of corporate brand directly. Corporate brand can concentrate and abstract corporate vision, culture, value view, behavior and the expectation of stakeholders, etc. as a kind of contract relationship through certain strategic action so that corporation and stakeholders can establish stable 'emotion connection'. For many emerging enterprises, in order to obtain competitive advantage through corporate brand management in growing period, strategic actions with great influence on corporate brand must be adopted.

3 Research Design

Based on the current corporate brand theoretical achievement, this research adopts cross-case comparison research method to carry out exploratory research on brand management method of mature enterprises. After analyzing corporate brand management activities for mature enterprises with different business scope, the development circumstances of mature enterprises' corporate brand presented with related strategic actions can be drawn, and then exploring ways of mature enterprises' corporate brand.

3.1 Sample Selection

To make sure the representativeness and comparability of selected sample, sample enterprises should be in different industries with differentiated products and management operation mode, as a result, the mature manufacturing enterprise and service-oriented enterprise are selected as samples for exploratory cross-case comparison study.

One case is Luzhou Laojiao, a mature manufacturing enterprise founded in 1982, which is a large brewing backbone group that developed from 36 ancient workshops in Ming and Qing dynasties, has long been history of working on the brewing of organic liquor and commits to advocate healthy lifestyle. Luzhou Laojiao received various prizes from its establishment, and its 'Luzhou' registered trademark is one of ten most famous trademarks in first Chinese trademarks election. Past brilliant achievement usually makes the company lose the direction, however Luzhou Laojiao, through successful corporate brand management and continuous strategic revolution, broke through the dilemma and returned as the leader in Chinese wine industry.

Another case is Jinjiang Hotel, a mature Service-oriented enterprise founded in 1958, which was assigned personally "Jinjiang" as the name by Marshal Zhu De and Marshal Chen Yi when it was completed, honored as the first five-star business and tourist hotel in Southwest China. The enterprise's age sometimes is the constraint affecting its development. Jinjiang hotel adopted series of strategic actions like strategic alliance to build and cultivate corporate brand when facing the dilemma. And brand of Jinjiang Hotel eventually becomes the industry leader in Southwestern China.

Basic situation of sample enterprises is as shown in Table 1.

Table 1. Basic information of sample enterprises

Name of enterprises	Luzhou Laojiao Co., Ltd.	Jinjiang Hotel Co., Ltd.
Time of foundation	In 1982	In 1958
Main business	Wine products	Hotel services
Nature of ownership	State-owned enterprises	State-owned enterprises
Scale of enterprise	Owns more than 5000 employees and total assets of 3 billion RMB, annual output of Luzhou series of liquor 50,000 tons, brand valuation is 345.84 billion RMB	A five-star hotel, belongs to the Sichuan Tourism Development Group, owns more than 700 rooms
Business scope	Mainland China	Mainland China

3.2 Data Collection

This research explores dynamic interaction of corporate strategic action and corporate brand through longitudinal case study. Main data collection starts from 2013, and ends in October 2015. During this period, various mixed data collection methods were adopted by starting from the purpose of explorative research, including Personnel interview, secondary data collection and on-site observation.

1. Personnel interview. Adopting half structural interview manner; interviews are enterprise management.
2. Documents collection. Collection object are sample enterprises' brief report, internal management handbook, product menu etc.
3. On-site observation. Observation objects are sample enterprises' headquarters, manufacturing base, sales and service places etc.

4 Case Analysis and Discussion

4.1 Case Analysis

(1) Luzhou Laojiao Co., Ltd.
In the 1990s, rapid development of China's economy and improvement of people's material life level stimulated the consumer's brand awareness. Especially in liquor-making industry, the high-end brand is more profitable. Maotai, Wuliangye, Jiannanchun have Stopped selling goods in nationwide and raised product prices to develop in the high-end market, while Luzhou Laojiao persisted in the road of mass brand and missed the development opportunities.

When Luzhou Laojiao realized failures in the high-end market, it tried to sale its high-end liquor products, but the brand unlimited expansion followed let Luzhou Laojiao face bigger difficulties. For making up previous wrong doings, Luzhou Laojiao adopted series of strategic actions to manage its corporate brand. It cut down 49 sub-brands and reserved core brand through sliming brand, cleaning up the market sorting out product structure. After 2 years, Luzhou Laojiao finished long-term brand development plans, and formulated reasonable brand strategy architecture including two brands as well as three series, shown in Fig. 1. It improved brand identification and consumer recognition. After that, to establish brand image and improve brand value, Luzhou Laojiao decided to stop selling goods nationwide and raise price. Simultaneously, Luzhou Laojiao started stock reforming and coordinated the relationship between the stakeholders. Additionally, Luzhou Laojiao integrated channel strengths through rewarding retailers and established brand image through actively participating social campaign. Luzhou Laojiao rebecame a leading enterprise in Chinese liquor industry with series of strategic planning and actions.

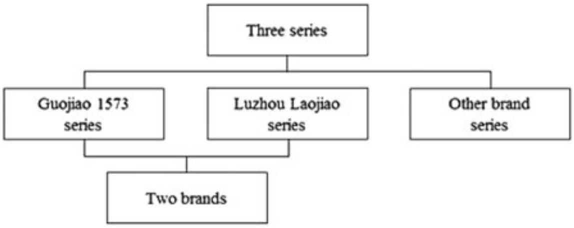


Fig. 1. Brand strategy architecture of Luzhou Laojiao

Corporate brand development process based on strategic actions for Luzhou Laojiao is as shown in Table 2.

Table 2. Corporate brand development process based on strategic actions for Luzhou Laojiao

Period and enterprise status	Important strategic actions affecting the corporate brand development	Corporate brand development situation
1980s Glorious history, high brand value and market acceptance	Establishment and development under the government's recognition and protection	1. Initially establishing corporate brand
		2. Strengthening brand value through old pits
1980s–1990s	Adhering to the mass brand Indulging in the past	Brand image was damaged
1997–2003 Inappropriate action under fear	1. Overstretched sub-brand	1. Brand excessive extension to destroy high-grade brand image of "Luzhou Laojiao"
	2. Taking the brand of "Luzhou Laojiao" as the intangible assets	2. Marketing model of paying money to OEM damaged the brand image
2002 - Brand renewal, value return	1. Sliming brand, cleaning up the market, sorting out product structure	1. Corporate brand achieving value return, Luzhou Laojiao embarking on the road of corporate brand management
	2. Stopping selling goods in nationwide	2. Rebuilding brand image in the minds of dealers and consumers
	3. Government approving listed companies to carry out tradable share reform	
	4. Non-public offering of no more than 3000 million shares	

(2) Jinjiang Hotel Co., Ltd.
1961–1979, as a hostel owned by government, Jinjiang Hotel didn't have authority or consciousness to manage its corporate brand. 1979–2001, because of the

reform and opening-up policy, Jinjiang Hotel went into modern enterprise administration phase of operating independence and self-financing. Since then, the time gave Jinjiang Hotel a great new identity. In 1995, Jinjiang Hotel, given the title of five-star hotel by National Tourism Administration, became the first five-star hotel in southeast of China. But with flourishing of Chengdu tourist market and increasing number of new hotels entering the market, fierce competition caused Jinjiang Hotel the development to receive the restriction again.

From 2001, the management team of Jinjiang Hotel decided to use advantage of carrying out the reshaping of the brand rebuilding work. It used "building a national top brand, being the best one of hotel industry" as vision, also offering services and products for high-end business market. And it redesigned brand logo, as well as inviting famous designer to redesign and redecoration the whole hotel. Then Jinjiang Hotel started building corporate culture and educational training system matched to corporate brand, and joined hands with strong brand enterprise. Through more than ten years of development, Jinjiang Hotel has successfully rebuilt its corporate brand through step-by-step strategic action.

Corporate brand development process based on strategic actions for Jinjiang Hotel is as shown in Table 3.

4.2 Case Discussion

According to the exploration of corporate brand management process regarding sample enterprises above, it can be found that the difference of business types can substantially make enterprises chose different strategic actions to build and manage corporate brand.

(1) Different extent of demand of corporate brand. Service-oriented enterprises possess more demand than manufacturing enterprises, since corporate brand is essential even only component of brand equity for service-oriented enterprises however, manufacturing enterprises have to concern product brand and related issues between product brand and corporate brand. Sample enterprises Jinjiang hotel relies on corporate brand to get involved in the competition, thus ways of adopting strategic actions to promote corporate brand are essential. Luzhou Laojiao has various product brand, it is significant to consider how to deal with the relationship between product brand and corporate brand through strategic actions.

(2) The difference of the essence of strategic choice. Service-oriented enterprises are keen on providing services with high quality to make success. However, manufacturing enterprises have to rely on innovation and R&D, high quality product and extraordinary manufacture to form strong corporate brand, which are implementation significance of strategic actions. Jinjiang hotel as sample enterprise has more concerns about services, and Luzhou Laojiao focuses more on product.

(3) Difference of implementation focus regarding organizational capability. Service-oriented enterprises tend to utilize intangible assets management to realize management goals of corporate brand. Manufacturing enterprises are suggested to improve tangible assets' operational capability to realize management

Table 3. Corporate brand development process based on strategic actions for Jinjiang Hotel

Period and enterprise status	Important strategic actions affecting development situation	Corporate brand the corporate brand development
1961–1979 Proud of "Yellow Vest"	Building and opening led by government	1. Initially establishing corporate brand
		2. Political overtonesbuilding strong brand image and basis
		3. Official hostel, lack of awareness of business operations
1979–2001 Develop-ment of new identity	1. Getting rid of government plans, becoming self-managed modern enterprise	1. Reforming corporate brand, laying the foundation of systematic brand management project
	2. Establishing enterprise management, system and reforming official human	2. Learning operation and management resource system
	3. Initiatingunion and establishing hotel management company	3. Becoming top1 hotel in southwestern China
	4. Carrying out hardware upgrading	4. Having some management infrastructure
	5. Regulating management	5. Difficult to resist external competition
	6. Awarded as the title of five-star hotel	
	7. Intense competition and inner defect	
2001 - Vigorous of brand develop-ment	1. Selecting brand development strategy	1. Carrying out series of ways to rebuild corporate brand
	2. Establishing vision, defining position, designing corporate logo	2. Discovering history heritage, increasing the essence of corporate brand
	3. Upgrading hardware facilities, improving product and service quality	3. Initially realizing the vision of "Nation's top and Industry leaders' brand"
	4. Embedding ideas through culture to prove service quality	4. Becoming the leading enterprise in Sichuan tourism and hotel industry
	5. Starting service personnel training system	5. Having ability to compete with international brands
	6. Focusing on opinions of leaders	
	7. Group expansion	
	8. Extending the brand network	
	9. Associating with strong brand enterprises	
	10. Rechanging management system	

goals of corporate brand. Sample enterprises Jinjiang hotel awakes consumers' memory through digging enterprises historical heritage and reforming image of corporate brand. Luzhou Laojiao improved the integrated value of corporate brand through re-positioning and forming product brand.

In the meantime, we found that from the points of strategic views, mature enterprises are suggested to consider tactics below to break through the dilemma of corporate brand management.

(1) Building the innovation expectation of corporate brand. And the construction of corporate brand has to be adjusted according to the changes of market conditions. At the first place, two sample enterprises immersed into the past distinctive achievement without improvement, which led corporate brand was lack of appropriate adjustment and revolution. Afterwards, in order to break through management delimma, Luzhou Laojiao conducted series of strategic actions to maintain its industry leading position.

(2) Knowing the methods of corporate brand system construction. Although sample enterprises were able to comprehend the significance of building corporate brand, but they still came across the crisis without adopted proper strategic actions. Luzhou Laojiao ignored the intrinsic connection between product brand and corporate brand in the process of exploring corporate brand management, so that the abuse of product brand weakened the value of corporate brand. However, through series of implementation, the corporate brand system was rebuilt. Therefore mature enterprises should clarify the position and role of corporate brand in the enterprise brand construction at the prior place. Designing the network structure of corporate brand, and adopting strategic actions, like putting the corporate brand as the core, to adjust brand structure is helpful for the strategy planning of corporate brand construction, which can the relationship and influence manner between corporate brand and product brand.

(3) Exploring the path of corporate brand rule development. Firstly, enterprises in different period of lifecycle, have different conditions of resources and capability and face different strategic choice, therefore management ways suiting corporate brand are different. And the difference of enterprise scale, the competition position, business type and operation ways can lead changes of corporate brand management ways. Enterprises are considered to constantly adjust corporate brand development path in accordance with development phases and strategic planning.

5 Conclusions

The research is based on strategic views, and integrates the strategic management and brand management theory. Besides, the research adopts the methods of vertical case study, and analyses ways of building corporate brand between mature manufacturing enterprises and service-oriented enterprises. From the process of building corporate brand for sample enterprises, implementing series of strategic actions which benefit prospects of future corporate brand development circumstance and direction, can boost the corporate brand construction of mature enterprises. Moreover, mature enterprises can adopt certain tactics to break through the dilemma of corporate brand management, thus realizing the corporate brand construction. It can be concluded that corporate brand construction and enterprises' strategic actions are correlated, mature enterprises brand construction is the consequence of related strategic actions, in the contrast, corporate brand construction has the certain impact on enterprises' strategic actions. The further discussion is placed regarding detailed roles, influence ways and path on enterprises' strategic actions from corporate brand.

Acknowledgements. We thank the support of the Projects of the National Social Science Foundation of China (16AGL003), Social Science Planning Annual Project of Sichuan Province (SC16C011), Soft science research plan project of Sichuan province (2017ZR0098, 2015ZR0091), Soft science research project of Chengdu city (2015-RK00-00156-ZF), The Central University Basic Scientific Research Business Expenses Project of Sichuan University (sksz201603, skqy201623, skqy201209), Projects of the Center for systems science and enterprise development (xq15B11).

References

1. Aaker DA (2006) Brand portfolio strategy. Strateg Dir 22(10):468–468
2. Armstrong G (2009) Marketing: an introduction. Pearson Education, New Jersey
3. Balmer J (2017) Advances in corporate brand, corporate heritage, corporate identity and corporate marketing scholarship. Emerald
4. Balmer JMT (2012) Strategic corporate brand alignment: perspectives from identity based views of corporate brands. Eur J Mark 46(7/8):1064–1092
5. Balmer JMT, Abratt R, Kleyn N (2016) Corporate brands and corporate marketing: emerging trends in the big five eco-system. J Brand Manag 23(1):3–7
6. Balmer JMT, Powell SM et al (2016) Advances in corporate branding. Palgrave Macmillan, UK
7. Buil I, Catalán S, Martínez E (2016) The importance of corporate brand identity in business management: an application to the uk banking sector. Bus Res Q 19(1):3–12
8. Chang A, Chiang HH, Han TS (2015) Investigating the dual-route effects of corporate branding on brand equity. Asia Pac Manag Rev 5(3):120–129
9. Dinnie K (2009) Taking brand initiative: how companies can align strategy, culture, and identity through corporate branding. J Brand Manag 16(7):496–498
10. Harris F, Chernatony LD (2001) Corporate branding and corporate brand performance. Eur J Mark 35(3/4):441–456
11. Hatch MJ, Schultz M (2001) Are the strategic stars aligned for your corporate brand? Harvard Bus Rev 79(2):128
12. Knox S, Bickerton D (2003) The six conventions of corporate branding. Eur J Mark 37(7/8):998–1016
13. Lewis S (2000) Let's get this in perspective. Unpublished presentation given at the Confederation of British Industry, Branding and Brand Identity Seminar, Bradford University School of Management
14. Mukherjee A, Balmer JMT (2008) Preface: New frontiers and perspectives in corporate brand management: in search of a theory. Int Stud Manag Organ 37(4):3–20
15. Srivastava K, Sharma NK (2013) Service quality, corporate brand image, and switching behavior: the mediating role of customer satisfaction and repurchase intention. Serv Mark Q 34(34):274–291
16. Tu YT, Li ML, Chih HC (2013) An empirical study of corporate brand image, customer perceived value and satisfaction on loyalty in shoe industry. J Econ Behav Stud 5:469
17. Voss KE, Mohan M (2016) Corporate brand effects in brand alliances. J Bus Res 69(10):4177–4184

Case Study: Packing and Distribution Logistics Optimization of Fashion Goods

Young Jae Jang(✉) and Shin Woong Sung

Industrial and Systems Engineering, KAIST, Daejeon, Korea
yjang@kaist.ac.kr

Abstract. In this short paper, we describe the development and implementation of the optimal packing and distribution of fashion products applied to a leading outdoor fashion brand in South Korea, Kolon Sport (K/S), through the academic-industry Business Analytics project with KAIST. We explain the framework of Business Analytics defined for the project and how modeling and analysis were performed under the framework. The packing and distribution decision involves how to determine a set of different box configurations and how to distribute these boxes to stores to meet store demand. The packing and distribution problem was formulated as a mixed-integer linear program, and a solution approach was developed to efficiently solve the industry-size problem in a timely manner. The modeling and analysis were performed with SAS packages. In particular, the optimization model was implemented with SAS/OR. The proposed approach was validated with simulation and on-site pilot test. The simulation showed an improvement in the packing and distribution process of around 15% in terms of meeting the demand in the stores, and it projected a $5 \sim 10\%$ improvement in revenue. After the successful pilot test, the approach was implemented in the internal IT system of K/S in July 2015. Consequently, the mass-produced items for the fall/winter season of 2015 were distributed by the system.

1 Introduction

1.1 Kolon Sport

Kolon Sport (K/S) is the signature brand of Kolon Industries, Inc., which was founded in 1957 in South Korea and operates more than 10 different fashion brands under its business unit. The total revenue of the Kolon Industries was around US $5 billion in 2014 [3]. The K/S brand was first established in 1973 and is now one of the top three outdoor fashion brands in Korea, a country with the world's second-largest market share in the global outdoor fashion industry, as of 2014, after the United States. K/S operates more than 250 stores over the country and designs, manufactures, and distributes more than 4,000 different types of products in all sales seasons. More information about the brand can be found at its website (http://us.kolonsport.com/). Note here that K/S refers to both the brand name and the department unit in Kolon Industries responsible for K/S operations. We use the term K/S interchangeably for the two definitions.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_133

1.2 Kolon-KAIST Business Analytics Project

The Kolon-KAIST Business Analytics project was first initiated in February 2014. The Big Data Analytics Team at Kolon Industries, an internal business analysis unit in the company that was established in 2013 to analyze massive amounts of data with advanced algorithms to support business operations and strategic decisions, was primarily responsible for the operation and management of the project. The team from the Korea Advanced Institute of Science and Technology (KAIST), the top science and technological university in South Korea, was responsible for guidance and for providing the technology for the project. The authors of this paper are the primary investigators of the project in the KAIST team.

The project goal was to improve the efficiency of the packing and distribution operation at K/S with Business Analytics. We defined Business Analytics as the process of delivering a specific tangible effect on the business operation by analyzing the data, developing scientific decision-making methods, and implementing methods with advanced IT solutions. Compared with conventional business analysis in which the final deliverables are forms of reports or documents of as-is analysis, the defined Business Analytics tried to provide a way to deliver specific actions and results. In contrast to the conventional business analysis or data-driven approach, which starts with collecting data and analyzing data first, we tried a problem-centric top-down approach. We first ask, "What is the problem?" and then ask, "Can this problem be solved with data and advanced algorithms and computation?" If the answer is yes to the preceding question, we try to collect data or create a system to collect the required data. Moreover, we emphasize an algorithm-based decision. To drive an action, "specific" action should be made. Therefore, optimization modeling and algorithm development are emphasized. We also include IT implementation in the definition of Business Analytics. To deliver sustainable results rather than a one-time as-is report, implementation of the analysis procedures and algorithm into the company's existing IT system is mandatory. This definition of Business Analytics clearly outlines the job scope of the project. It includes problem identification, data analysis to find the insight, development of an optimization model and algorithm to solve the complex problem, and then implementation of the system to derive the action.

Figure 1 shows the framework of Business Analytics that we constructed for the project. The framework describes the required techniques and corresponding departments or organization necessary to support the techniques. As shown, the framework consists of three different technical layers. The bottom layer comprises data collection and data management. The fact-based and data-oriented approach is the core concept of Business Analytics. Identifying the required data and effectively managing them were the fundamentals of the project. The IT department at K/S was responsible for this layer. The next layer addresses knowledge discovery through the analysis of data. Using data mining and statistical analysis, we tried to find new insights and knowledge or validate the existing knowledge. This function was done in cooperation with the

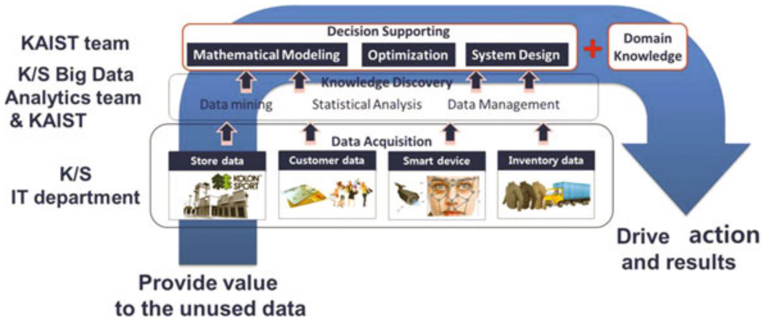


Fig. 1. Business analytics framework for the project

marketing team, business operation team, and various analytics-centric teams at K/S under the guidance of both the K/S Big Data Analytics and KAIST teams. The top layer concerns decision making. With the knowledge and insight gained from the underlying layers, this top layer attempts to support the decision for a complex problem. Mathematical optimization techniques were used in this layer. Because K/S did not have an internal team or organization using such optimization, this layer was the primary responsibility of the KAIST team.

2 Packing and Distribution Problem

To expedite the shipping process and reduce the operational costs of the distribution, K/S has been using direct delivery of their apparel products from the manufacturers, which are the third-party OEM supplier to K/S, to the K/S stores for the last several years. In the past before direct delivery was used, all of the manufactured products were shipped to the K/S warehouse, unpacked, sorted, and repacked for distribution to the stores. Given the fact that K/S distributes more than 1.6 million items in 1,000 different styles during each selling season, this repacking process required significant time and effort. Moreover, due to long delays in the repacking process, some stores did not receive products even after a sales season had already started. Direct delivery provided a significant benefit to the process. Each manufacturer packs the product with different sizes based on the size configurations given by K/S and ships their goods direct to each store. For example, K/S requests the manufacturers to provide two different types of box configuration such that for product WJ0001, Box 1 should contain 10 small, 30 medium, and 20 large sizes, and Box 2 should contain 3 small, 5 medium, and 4 large sizes. Also, K/S informs the manufacturer of which stores get how many boxes of each box configuration.

Although direct delivery significantly eliminated inefficiency in terms of delivery speed, it generated another problem. Because packing is done by third-party manufacturers, it was not realistic to make different box configurations for every store for each product, which would cause significant operational costs particularly when the numbers of products and stores are large. Therefore, the number of different box configurations was restricted to less than 10. Consequently, K/S should decide how to design the box configurations and determine which configuration and how many boxes are distributed to a store. These decisions are collectively called packing and distribution in K/S. Tables 1 and 2 show an example of packing and distribution decisions. The decisions on the configuration for each box type and the number of boxes for each type should be made in the packing and distribution process. In addition, which types and how many boxes of each type should be sent to each store must also be determined.

Table 1. An example of the packing decision

Configuration type	Configuration	Number of boxes
Type 1 Box	XS: 20, S: 25, M: 40, L: 15, XL: 10	10
Type 2 Box	XS: 10, S: 15, M: 20, L: 12, XL: 5	50
...	...	...
Type n Box	XS: 5, S: 10, M: 12, L: 8, XL: 3	30

Table 2. An example of the distribution decision

Store	1st box		...	m^{th} box	
	Configuration type	Number of boxes	...	Configuration type	Number of boxes
Store 1	Type 3 Box	1	...	Type 4 Box	1
Store 2	Type 4 Box	3	...	-	-
...	...	...	...	...	...
Store s	Type 2 Box	2	...	Type 3 Box	1

Because the OEM manufacturers charge packing fees based on the number of configuration types and the number of boxes shipped to the stores, appropriate packing and distribution can significantly reduce costs. However, due to the complexity of the configuration, decisions had been made on an ad-hoc basis that had caused inefficiency - some stores received more items than needed and some received less than needed.

3 Project Goal and Scope

The goal of the project was to develop a decision support system to optimally determine packing and distribution. The project was broken down into the following tasks:

- Demand analysis for each product style and size using the past sales data;
- Optimization modeling and algorithm development;
- Validation of the model and algorithm with simulation;
- Validation of the overall approach with pilot tests;

- Implementation of the approach in K/S's internal IT system.

The K/S Big Data team took the leading role of managing the overall project. The KAIST team took a leadership role for the first, second, and third tasks. The Sales and Operation Team at K/S was responsible for the fourth task, and the IT Department at K/S was responsible for the final task. The KAIST team also worked as the technical advisor for tasks four and five.

Note that demand forecasting or evaluation of the quantity of each product type was not within the scope of the project because of the following reason. The packing and distribution decisions were needed about one to two months before each sales season starts. However, decisions on production quantity had to be made one year before the target sales season. That is, by the time they were ready for packing and distribution, all of the products would have been ready for shipment at the manufacturers. Therefore, the product quantities for the items were given as the input parameters for the box assortment.

Although the product quantities to be distributed were known, the number of each item to be distributed to each store still needed to be evaluated. For instance, given that there are 100 medium-size winter jackets (item number WJ000111) available for the season, the number to distribute to each store must be determined. We evaluated this quantity based on the past sales data. Specifically, for each product style, we evaluated the percent of sales made by each store. This sales contribution value from the past data was used to estimate the distribution quantity for each product. For statistical analysis, we used multiple regression analysis and a standard method to estimate the future demand.

4 Packing and Distribution Optimization

The core part of the project was the development of the optimization model in the second task. As mentioned above, in this model, we determine the number and contents of the box configurations and their allocation to the stores to minimize the difference between store demand and the actual distribution quantity as depicted in Fig. 2. As indicated in the figure, the decision variables in the optimization model are the configuration of each box and the number of boxes to be distributed to each store. The demand for each store is evaluated from Task 1.

The constraints for the model are as follows. First, the types of configuration are limited. The manufacturers do not want to have more than 7 different types of configurations for each product. Second, the maximum number of boxes is restricted, as is the capacity of the box. Of course, the maximum capacity depends on the type of product. A $50 \times 50 \times 50\,\text{cm}^3$ box can hold 10 puffy winter down jackets at maximum, whereas it can hold 50 summer T-shirts. K/S has data mapping on the size of each box type and its maximum capacity for each product type. In addition, the maximum and minimum number of boxes a store receives are constrained. The last constraint is the upper limit of overstocking for each store. Note that due to the discrete nature of the problem, it may be impossible to meet the exact demand for each store. Therefore, some levels of over-shipping or under-shipping are unavoidable. However, K/S is more concerned with over-shipping. As a result, we added an over-shipping constraint, which specifies the upper limit of the allowable over-shipping to each store. The details of the general mathematical model are provided in the Appendix. The solution algorithm was also developed but is not presented in this document. For readers interested in the specifics of the algorithm, please contact the corresponding author.

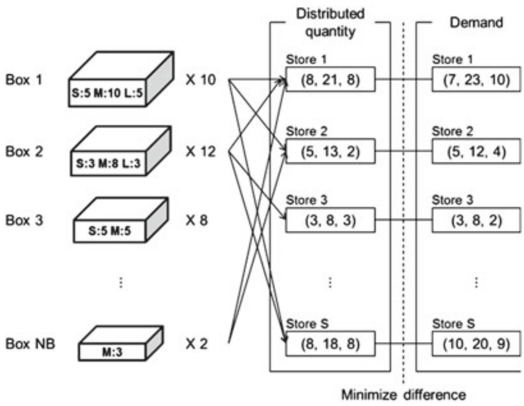


Fig. 2. Concept of the box assortment and allocation optimization problem

Similar formulations and optimization models are available in the academic literature, including [1,2]. However, the details of the models and the scales of the problem these optimization problems are trying to solve are different. To the best knowledge of the authors, this is the first work to develop an optimization model and solution algorithm at an industry-level scale and apply them to an actual business operation. Note that this paper is not intended for the academic community, and therefore the academic value of the work and scholarly implications are not discussed.

5 Validations

We conducted a simulation-based validation of the packing and distribution process in one sales season to compare the performance of the proposed method to that of the legacy procedure at K/S. The sales volume, i.e., number of sales, from the proposed method was approximately 10.16% better than that with the legacy method. We also conducted a pilot test of our approach by distributing a number of actual products for the summer season of 2015, and the result shows that compared to the legacy method of K/S, the inventory was distributed well by our method in accordance with store demand. Finally, the proposed method has now been completely implemented in the K/S's internal system for use in the upcoming packing and distribution season. The details of the validation results can be found in [5].

6 Conclusion

In this document, we describe the Business Analytics project jointly conducted by K/S and KAIST. The project has successfully developed a decision support system to assist the packing and distribution procedure at K/S. Throughout the project, the framework of the Business Analytics was first defined, and the roles and responsibilities of each team were clearly identified based on the framework. Then, multiple tasks were listed and performed by the joint team. The proposed packing and distribution approach includes a statistical data analysis that estimates the demand of stores for each product, the optimization of the box configurations and their distribution, and the IT implementation for sustainable use. The core technologies developed in the project were the optimization modeling and algorithm development. The box configuration of each box and the distribution quantity for the boxes to the stores were optimized by mixed integer programming. The proposed method was validated with simulation and actual tests in the stores. As a result, the proposed method showed significant effect, and K/S decided to implement it within their internal IT system.

The financial benefit delivered by the project is still under investigation. Based on the simulation result and on-site tests, we project that the proposed packing and distribution will improve sales by 5% 10% in terms of revenue. From the cost-saving viewpoint, we estimate that costs related to the man-hour working time for determining the configuration, the boxes, and shipping and handling will be significantly reduced. The unquantifiable contribution of this project is also very significant. The project demonstrates how advanced data-driven methods and algorithms can be incorporated into a traditional retail fashion business. It proves that in the modern data-driven society, a high-technology company is not defined by its product but rather by how it handles its operations. Mr. J.H. Jang, the Lead Manager of the Big Data Analytics Team, stated that "this project showed how the scientific method can improve the operation and provided the direction of the Big Data Analytics Team".

Moreover, this project has been an example of successful academic-industry collaboration deriving actual tangible results. Often, academic researchers find topics for research from other academic literature or come up with hypothetical imaginary topics in the hope that these kinds of problems will be valuable to industry. However, the KAIST researchers identify multiple topics that are not known to academia but which are worthy of further investigation. The authors are also currently working on an academic paper presenting the optimization algorithm they developed for this project.

Finally, this packing and distribution problem is common across the retail fashion industry. To our best knowledge, few retail fashion companies use logical methods for the process. The process, model, and algorithm developed in this project can be further developed as a service or software solution. We discovered three patents (two applications and one awarded patent) for solution algorithms of the optimization problem similar to what we constructed [4,6,7]. Among these patents, two are from the Oracle Corporation [6,7] and one is from the SAS

Institute [4]. However, we found that there is still room for improvement in the algorithms presented in these patents.

Appendix

Box assortment optimization model

The optimization model determines the number and contents of the box configurations and their allocation to the stores to minimize overstocking and understocking. The demand for items in each store is given by a multiple regression model. Because we know the maximum demand for each item among all stores and the capacity of the box is finite, it is theoretically feasible to generate a pool of all possible box configurations by selecting numbers between zero and the maximum demand for each item and making a combination of them as a box configuration, only if the sum of the number of items in a combination satisfies the capacity. In this way, a set of possible box configuration is given.

When the box configurations and their assignments are established in K/S, several managerial and operational requirements need to be met. First, a box only contains items with an identical product type, style, and color, but different sizes. For example, for such an item with five different sizes, a box can contain two "Small"-sized items, five "Medium"-sized items, and three "Large"-sized items because the manufacturer prefers to pack items of the same product type, style, and color at one time unless the finished products are waiting until changes in production setup and raw materials are needed. Additionally, the manufacturers of the items might be different even if their product types and styles are the same but their colors are different. Second, the following are constrained: the number of different box configurations, the maximum and minimum quantities of items in a box, i.e. capacity, the maximum and minimum number of boxes a store receives, the maximum number of boxes, and the upper bound on overstocking.

Indices

S : set of stores ($s \in S$);
I : set of items ($i \in I$);
C : set of possible box configurations ($c \in C$)

Parameters

d_{is} : demand of item i in store s;
c_{bi} : integer number of item i in box configuration b;
α_{is} : understocking penalty of item i in store s;
β_{is} : overstocking penalty of item i in store s;
δ_{is} : upper bound on overstocking of item i in store s;
NB : number of different box configurations;
N_s : minimum number of boxes store s receives;
M_s : maximum number of boxes store s receives;
T : maximum number of boxes

Decision Variables
x_{bs} : number of box configuration b allocated to store s;
y_b : binary variable, which is 1 if box configuration b is used and 0 otherwise;
u_{is} : understocking of item i in store s;
o_{is} : overstocking of item i in store s

Objective
$$\sum_{s\in S}\sum_{i\in I}(\alpha_{is}u_{is}+\beta_{is}o_{is})$$
Constraints
$$\sum_{b\in B} c_{bi}x_{bs}-o_{is}+u_{is}=d_{is}\quad \forall i\in I, s\in S$$
$$\sum_{b\in B} x_{bs}\geq N_s\quad \forall s\in S$$
$$\sum_{b\in B} x_{bs}\leq M_s\quad \forall s\in S$$
$$\sum_{b\in B} y_b\leq NB$$
$$\sum_{b\in B}\sum_{s\in S} x_{bs}\leq T$$
$$\sum_{s\in S} x_{bs}\leq \sum_{s\in S} M_s y_b\quad \forall b\in B$$
$$\sum_{s\in S} x_{bs}\geq y_b\quad \forall b\in B$$
$$x_{bs}\geq 0, \text{integer}\quad \forall b\in B, s\in S$$
$$y_b\in\{0,1\}\quad \forall b\in B.$$

References

1. Fischetti M, Monaci M, Salvagnin D (2015) Mixed-integer linear programming heuristics for the prepack optimization problem. Discrete Optim 22:195–205
2. Hoskins M, Masson R et al (2014) The PrePack optimization problem. Springer
3. Industries K (2014) Financial information http://www.kolonindustries.com/Eng/Service/service02.asp
4. Pratt RW (2014) Computer-implemented systems and methods for pack optimization. U.S. Patent Application
5. Sung SW, Jang YJ (2015) Kolon-KAIST business analytics project. Internal report, Kolon-KAIST Life Innovation Center
6. Vakhutinsky A (2012) Retail Pre-Pack optimization system. U.S. Patent Application
7. Vakhutinsky A, Subramanian S et al (2012) Retail pre-pack optimizer. U.S. Patent Application

Integrated Project Management

Algorithmical and Program Functions of Innovation Project Management in Technoloji Park

Elchan Ghuseynov, Javanshir Mammadov, and Shafahat Rahim Rahimov(✉)

Sumgait State University, Sumqayit, Azerbaijan
shafahat_61@mail.ru

Abstract. It was analyzed the present status of the innovation projects management problem in many technology parks of European, Asian and American countries. On the base of some of exists methods of option and creation process of innovation projects, there were defined the basis directions by innovation project management in technology park, also scientific problems of subject area and option of the best projects by means of flexible complex program system. In depend on the defined scientific problems in the paper, the process of managing an innovation project on the beginning stage of working in the departments of the technology park at higher educational school in Azerbaijan was considered. On the basis of the priority scientific directions and the scientific profiles of scientific technology park in the Sumgait State University (SSU) have been defined ways of decision a problem of an innovation project management from its registration till its option and input to the data base by an expert. On the base of the structural scheme of technology park in SSU, its scientific profiles, infrastructure and specialists, a new conception managing innovation projects in the technology park of SSU in Azerbaijan by means of complex flexible interface was offered. For executing all procedures of innovation project management on the stages its creation, experimental research, manufacture and business, special algorithm was worked out. By means of this algorithm, the block-scheme of managing functions, the program tool for development of new innovation project in the technology park beginning from the registration of the project till its assessment was created.

Keywords: Technology park · Flexible manufacture · Innovation project · Management · Higher educational school · Project assessment

1 Introduction

As known, creation of new innovation projects in the many science technology parks, Zouain [13] and Lamperti [7] considered the problems of science technology parks from its laboratory investigation, manufacture till economical process

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_134

in Brasilia organization is complex, difficult scientific-research, engineering, manufacturing and economical managing process. Traditional technology parks in of European, Asian and American countries solve own problem of innovation project management by means of the corporative connections between different departments of the technology park Volkonitskaia [12], Kirchberger and Pohl [6] considered the ways of decision of functions innovation project development in technology parks of different countries "without application of special software systems what do not enough provide efficiency of planning process of creation of innovation project. Application of such program will allow to realize some science-technical, collection of necessary information about the new project, its storage, choosing, expert analyzing and assessment. Analyze of the existed methods and ways of new innovation projects development Murzina [8] and Díez-Vial [4] considered international practice of innovation structure creation with economical knowledge for technology parks" shown that, there is not identify methodology for optimal choosing innovation project, its algorithmically and program providing, which would be allowed to support high economical efficiency and get great financial results for the budget of technology park.

Analyze of the problem by innovation projects option for them scientific research, manufacture and commercial process shown that the date problem solved without complex system application on a level of some experts solution after presentation of projects. In this connection, for effective solution of the best innovation project option under the complex system there were determined problems of researching by the following directions:

- Determination of the basic scientific profiles in the technology park at Sumgait State University (SSU) which can provide option of the necessary innovation projects for development them in the parts of science, manufacture and business of technology park;
- Creation of new conception of managing innovation projects in the technology park of SSU in Azerbaijan which can provide realization of option the best innovation project in the complex system for efficiency development of technology park for short time;
- Development of algorithmically and program tools for providing management of innovation projects in the scientific technology park of SSU which can provide executive menu procedures in corporative network system of technology park "Rafael Avila Faraco [9] considered a problem of creation and using social networks and knowledge transfer in technological park companies in Brazil, what is solved only forming data base in the internal system and therefore procedure of the best innovation project is done intuitive".

In depended on the determined scientific problems, the purpose of the paper defined as follows: it is necessary to work out universal algorithm and program which would be allowed to create innovation projects in technology parks with different scientific profiles.

This material was written by means of the executed sciences researches at the department of "Information technology and programming" of Sumgait State University, Azerbaijan Republic.

2 Determination of the Basic Scientific Profiles in the Technology Park at SSU

At present in Azerbaijan Republic, different innovations programs in many application areas are realized. Specially, in many high educational schools of Azerbaijan, as product of innovation technology, new technology parks are created. Traditionally at first stage of technology park creation, a basic problem is option of the necessary scientific profiles in the universities, as they have different specifies of education and different balance of scientific potentials-experts "Correia and Gomes [3] considered a problem of forming potentialities and limits for the local economic and innovative development of technology parks located in the Northeast region of Brazil, where the problems of scientific directions are limited owing to little application of intellectual information systems in different areas".

As an object of investigation is SSU, where scientific technology park must be created, then considering specifics of specialties of this universities and more development scientific directions of SSU, the basics profiles of technology park had been chosen:

- Application of intelligence information systems in different area of sciences (practically in medicine, biology, physic and other areas);
- Investigation of physical and chemical process and theirs application in manufacture;
- Application of modern commercial methods and systems for efficiency realization of technology park.

The first scientific profile of SSU technology park is one of priorities directions in technology park, as application of progressive information technology, intelligence systems and program tools provide computing functions of many medical, biology, engineering and others important technical problems for citizen. Option of this scientific profile was justified by means of the formed scientific base which organized scientists of SSU, their scientific works which they conduct about 50 years. Under scientific leader of Rafik Aliyev, a famous in world scientific circles by development of fuzzy logic theory, was organized the scientific school of scientists of SSU and theirs works had been implemented in flexible manufacture systems in the metallurgy enterprise of Sumgait city. At present the scientists as doctorates and professors work in SSU by own specialty. At present 4 the laboratories of the departments of faculty "Engineering" provide all education and sciences works of students and teachers.

The second scientific profile of SSU scientific technology park is multi-scientific area which is connected theoretical and experimental investigations, also implementation of results in manufacture process. Investigation on this directions conduct scientists of physical and chemical faculties which have 4 the scientific laboratories and some special devices for experiments, are founded in the scientific part of SSU.

The third scientific profile of SSU scientific technology park (SSU STP) is one of the basis profile, as that profile must provide economical realization of two profiles of SSU scientific technology park in whole. Application of modern commercial methods and systems of management and marketing provide all way economical control of innovation project beginning from an idea appearance, its including to the technology park system, presentation, theoretical and practical researching in laboratory, optimal option, embedding in the flexible manufacture and sale an innovation product in the internal or external market.

On the base of the explained scientific profiles of the technology park, we can mark that the existed laboratories for experimental researching in the first and second profiles are science C research part of the technology park, and in future they will have to be able to work by connection with business incubator and flexible manufacture.

3 Managing Innovation Projects in the Technology Park of SSU

Efficiency of working the experimental laboratories by 3 scientific profiles depends on conception of managing innovation projects in the technology park, as some of this projects have not the necessary text, graphical, animation and video information for expert analyze, high technical and economical characteristics, and also possibility of them making in the flexible manufacture system owing to lack of some equipments and materials "Prencipe [10] considered a problem of realization of innovation process in University Spin-offs" what is shown only some innovation operation by economical management and marketing. In this connection, solution of the problem of new conception creation of managing innovation projects in the technology park of SSU in Azerbaijan is important, scientific-actuality researching problem.

In accordance with the offered conception of managing innovation project in technology park, using principals of flexibleness and openness of working the automation system, the way of managing innovation project is represented by means of the following scheme:

- Input and presentation of scientist user idea with the necessary data of own user;
- Expert analyze of project annotation;
- Experimental investigation in the laboratory and making a laboratory exemplar of the project;
- The project embedding in flexible manufacture and its output Elchan Ghuseynov, Javanshir Mammadov, Gulnara Genjeliyeva [5] considered a problem of application of flexible industrial park in the scientific technology park of Sumgait State University of Azerbaijan where the problems of innovation project option and embedding in the manufacture are not considered.

4 Algorithmically and Program Tools for Providing Management of Innovation Projects in the Technology Park of SSU

Functions of management departments of SSU STP and development, commercial way of innovation project are provided by means of the algorithm scheme shown in the Fig. 1. The Procedures of innovation project management from the beginning registration to its embedding into manufacture are represented on the base of the following stages:

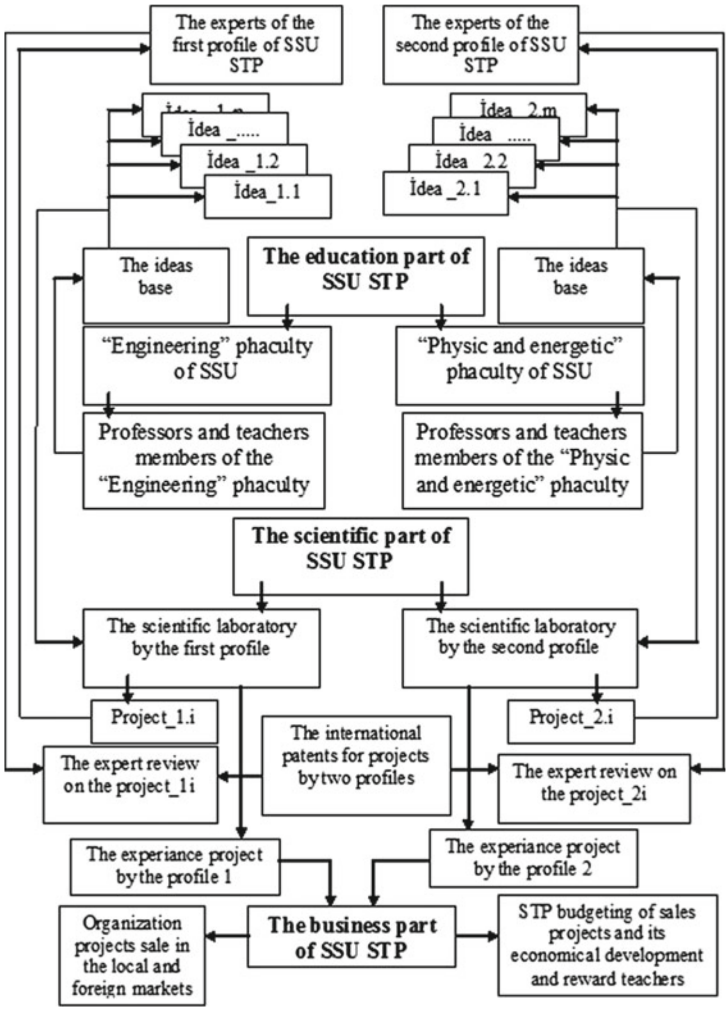


Fig. 1. The structural scheme of SSU STP with innovation project relocation

(1) The professors and teachers, also students of SSU present own ideas in the global level of the system of registration. At first they input the registration information to two the panels “Teacher registration” or “Student registration”. An User of the system inputs his data and information about himself project (name, purpose of the project, short idea of the project-annotation) which is saved in the “Annotation” panel for future familiarization of this project by an expert. Also on the first stage for careful analysis in this part of the system, the panel of “Prototipe of a project” is used.
(2) During a week, different experts in autonomy form control the registration and annotation informations. On the base of standart form of the new project receiving, a decision about acceptance or rejection of the project are given. Procedures of an expert are executed and saved in the panel “Expert review”.
(3) For carefully presentation of the projects, the projects are represented to the experts in the determined time. Actuality, newness, modernity, high engineering decision and economical efficiency of the projects are the basis of choosing the project for its first making as an experimental form in the laboratory condition in accordance of the scientific profiles.
(4) In the experimental laboratory the best project is given for its constructor design, material option, working out the first experience form of the project and control of its mechanical, automation and others functions. At this stage the basics technology characteristics of the project are defined and then its quality level, the better data in difference of the prototypes ones. The documents must be official formed by the status leader of the laboratory.
(5) All information of the designer is saved in his data base (in the part of “Project data base”).
(6) The commercial department determines the basis of rules and the demanded in the external and internal market and economical efficiency of the project is computed. Čirjevskis [11] presented dynamically “signature business model” which can provide durable competitive advantage at commercial procedures of business incubator of technology park. “For presentation of the project to the local and international markets Amit and Zott [1] worked out the model for application in business process of innovation project by science and technology connection” in “Presentation of the new project” department, the view of the 2, 3-measure pictures, animations, video and technical characteristics are saved in the data base. The managers by the scientific profiles chooses the clients and some information about them saves in the “Client firm”. Between the scientific technology park and the firm, official meeting is executed by means of this department. Therefore the official documents are prepared.
(7) The procedures of project documents presentation are realized for the experts in reality. On that stage all the documents registered by the expert are sent to the flexible manufacture where a process of the project making is executed.

The structural scheme of SSU STP with innovation project management is given in Fig. 1. Also the functional connections between the jobs departments of SSU STP are shown.

The represented text and block-scheme algorithms allow to develop a program for checking, choosing a project way in the scientific technology park at the Sumgait State University. In this connection, structural scheme of an innovation project management is proffered on the base of this algorithms (Fig. 2). The offered scheme consists of three blocks where each one works in different rooms, beginning from the room of projects registration, the room of the saved data base and the experts center.

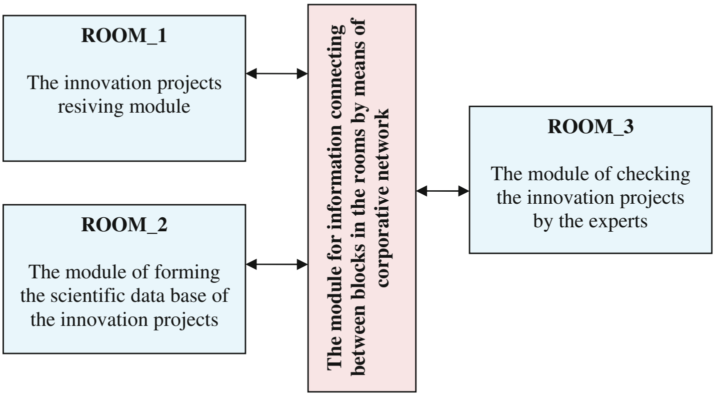


Fig. 2. Structural scheme of an innovation project management under corporative network of technology park

The innovation projects receiving module provides presentation of the basis data about authors, title and aim of a project, receiving date and time of project in the system, annotations of the project and its figure. The module of forming the scientific data base of the innovation projects provides automatically saving all data about a innovation project in the data base. Also in this database like table form, the certification by application area is executed. For assessment of every project which was marked in the system, an expert by corresponding application area checks his innovation project and gives a mark by the development project.

Checking one project can exist by some experts in this application area. Information about assessment of a project is sent to an author of the project. In this message answer can be by some versions: accepted; accepted, but must be corrected; rejected.

5 Development of Program Tool for Expert Assessment of Innovation Project in Scientific-Technoloji Park

In corresponding to functions of each the module, the program for expert assessment of innovation project (EAIP) Carver et al. [2] considered the problem of development of special Issue on software engineering technology and its applications corporative system of manufacture what is shown some program procedures of designing innovation project in corporative network is worked out (Fig. 3). The program for EAIP is worked by the following stages:

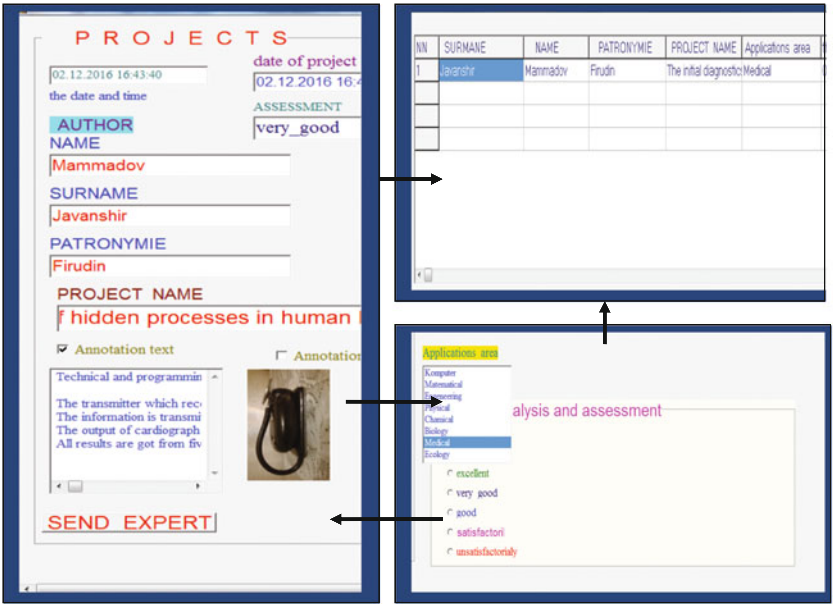


Fig. 3. The scheme of program functions in the EAIP

(1) At first stage a designer inputs his the data of authors names, title and aim of a project, annotations of the project, the reciving date and time of project into the system.
(2) On second stage all the data by the project are saved in the data base management system as a table view.
(3) In third stage an expert checks information about the designer of a project and begins to read and looking annotation of the project. After checking the project, an expert assessment of this project gives by marks where a result outputs in the system of EAIP and sends to the designer.

6 Conclusions

On the base of the made investigation by innovation project management, the following results were got:

(1) In corresponding to the aim of the investigation problem, the algorithm procedure for management and assessment of an innovation project in scientific-technopark was offered.

(2) The block-scheme of innovation management project, forming data base of projects and assessment of a project by an expert was worked out.
(3) On the base of algorithm procedures of innovation project management, the program operations with project registration; its author and text names; text and graphical annotations; sending all information to expert by different scientific areas; saving the neccessary data in the data base; expert assessment of innovation project were developed.

References

1. Amit R, Zott C (2014) Business model design: a dynamic capability perspective. Technical report. Citeseer
2. Carver D, Chan WK et al (2016) Special issue on software engineering technology and applications. J Syst Softw 126:85–86
3. Correia AMM, Gomes MDLB (2014) Potentialities and limits for the local economic and innovative development: a comparative analysis of technology parks located in the northeast region of Brazil. Int J Innov Learn 15(3):274–298
4. Díez-Vial I, Fernández-Olmos M (2015) Knowledge spillovers in science and technology parks: how can firms benefit most? J Technol Transf 40(1):1–15
5. Ghuseynov E, Mammadov J, Genjeliyeva G (2016) Application of flexible industrial park in the scientific technology park of Sumgait state university of Azerbaijan. Br J Appl Sci Technol 13(4):17–26
6. Kirchberger MA, Pohl L (2016) Technology commercialization: a literature review of success factors and antecedents across different contexts. J Technol Transf 41(5):1–36
7. Lamperti F, Mavilia R, Castellini S (2015) The role of science parks: a puzzle of growth, innovation and R&D investments
8. Murzina SE (2015) International practice of innovation infrastructure creation as a mechanism for the innovative economy development and the improvement of land use effectiveness. J Siberian Federal Univ Humanit Soc Sci 8:2535–2544
9. Mussi C, Angeloni MT, Faraco RA (2014) Social networks and knowledge transfer in technological park companies in Brazil. J Technol Manage Innov 9(2):172–186
10. Prencipe A (2016) Board composition and innovation in university spin-offs: evidence from the italian context. J Technol Manage Innov 11(3):33–39
11. Čirjevskis A (2016) Designing dynamically "signature business model" that support durable competitive advantage. J Open Innov Technol Mark Complexity 2(1):15
12. Volkonitskaia K (2015) Business models of technoparks in russia. In: Series: Science, Technology and Innovation WP BRP, WP BRP 55/STI/2015
13. Zouain DM, Plonski GA (2015) Science and technology parks: laboratories of innovation for urban development - an approach from Brazil. Triple Helix 2(1):1–22

Two-Stage Fuzzy DEA Models with Undesirable Outputs for Banking System

Xiaoyang Zhou[1,2], Rui Luo[1], Benjamin Lev[2], and Yan Tu[3](✉)

[1] International Business School, Shaanxi Normal University, Xi'an 710119, People's Republic of China
[2] LeBow College of Business, Drexel University, Philadelphia, PA 19104, USA
[3] School of Management, Wuhan University of Technology, Wuhan 430070, People's Republic of China
tuyan_belle@163.com

Abstract. In this paper, we propose two stage fuzzy DEA models with undesirable outputs to evaluate banking system. The banking system is divided to two subsystems: production stage and profit stage. In the model, two kinds of assumptions (constant returns to scale and variable returns to scale) are considered, and fuzzy parameters are adopted to describe the uncertain factors. Chance constrained operator is used to handle the proposed model and equivalent transformations are given to make the models solvable. We illustrate and validate the proposed models by evaluating 16 Chinese commercial banks. Some discussions are also presented to show the differences and advantages of the models.

Keywords: Two-stage DEA · Banking system · Undesirable outputs · Constant returns to scale · Variable returns to scale

1 Introduction

Efficiency is particularly illustrated in the comprehensive competitiveness of commercial banks. In terms of investment, supervision or independent operation, it is of great significance to evaluate the efficiency of commercial banks. Data Envelopment Analysis (DEA), a non-parametric mathematical programming approach to evaluate group of Decision Making Units (DMUs) with comparative efficiency, is widely applied to analyze the efficiency in the process of banking operation [1,4,5,12,13,17].

It seems like the traditional DEA model treats DMU as "black box", which is unable to explain the internal structure inside the "black box". In a banking system, taking deposits and making loans are the main activities. Some banks may have greater advantage in taking deposits, but perform not as well in making loans to realize profit. Hence, it is better to consider the banking system as a two-stage process. We should evaluate not only each stage, but also the whole system.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_135

In the banking system, there are desirable outputs, like the total deposit during the first stage, and the profit during the second stage. But in the operational process, banks want to get rid of the risk exposure and non-performing loans/assets, which are undesirable outputs. Scheele [15] concluded that the undesirable outputs could be handled by direct and indirect approaches. In indirect approaches, the values of the undesirable outputs are converted to a monotone decreasing function, and then we can treat the undesirable outputs as the other normal desirable outputs. Direct approaches avoid data transformation and incorporate the undesirable outputs as the inputs directly into the DEA models. In Hu's work [20], undesirable outputs were treated as the inputs, and the less of the undesirable outputs (inputs), the better. A new two-stage DEA model with the undesirable outputs to measure the slacks-based efficiency of Chinese commercial banks during years 2008–2012 was developed in [2], where the banking operation process of each bank was divided into a deposit generation stage and a deposit utilization stage. However, they neglected the undesirable outputs of the deposit generation stage.

In recent years, scholars have already realized that some inputs and outputs are imprecise in real life cases. Conventional DEA models used the accurate inputs and outputs data, which may not fit in the real world cases. To deal with the imprecise data, the fuzzy theory was introduced in DEA to become fuzzy DEA.

According to Hatami-Marbini [7], the methods of handling the fuzzy DEA can be divided into four types: (i) The tolerance approach [16]; (ii) The α-level based approach [8,11]; (iii) The fuzzy ranking approach [6]; and (iv) The possibility approach [9,10,19]. Puri [14] presented a fuzzy DEA model with undesirable fuzzy outputs, giving a numerical illustration of the banking sector in India using fuzzy input/output data for the period 2009–2011, but the internal structure of banking operation system was not taken into account. Wanke [18] developed new Fuzzy-DEA models to measure the impact of each model on the efficiency scores and to identify the most relevant contextual variables of efficiency by using bootstrap truncated regressions with fixed factors.

In this paper, we propose a two stage fuzzy DEA model with undesirable outputs to evaluate 16 China's commercial banks. The remainder of this study is organized as follows. Section 2 presents problem statement. Section 3 develops three kinds of two-stage DEA models with undesirable outputs. Section 4 discusses the solution method. Section 5 gives a case study with 16 banks in China. Finally, conclusions are given in the last section.

2 Problem Statement

The efficiencies of several banking systems (DMUs) are under evaluation, and each DMU, indexed by $j = 1, \cdots, J$, has two sub-stages: production stage and profit stage. During the production stage, funds are collected from depositors while consuming resources such as number of employees x_j^{NE}, fixed assets x_j^{FA}, operating expenses x_j^{OE} and number of institutions x_j^{NI}. Total deposits Z_j^{TD} is not only the desirable output of the production stage but also the input to

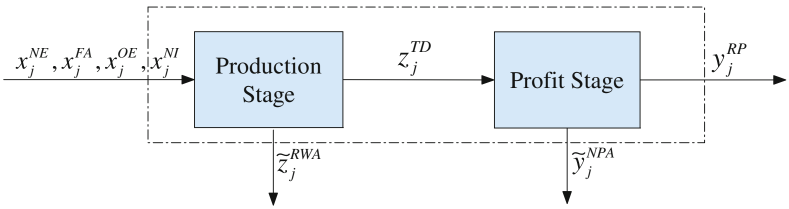


Fig. 1. Two-stage banking system

the profit stage. Total risk-weighted asset $\tilde{z}_j^{\mathrm{RWA}}$ is the undesirable output during the production stage. In the profit stage, deposits Z_j^{TD} are used to invest into other activities to get profits y_j^{RP}, but the undesirable output non-performing loans/assets $\tilde{y}_j^{\mathrm{NPA}}$ will also be produced. Because deposits collected during the production stage determines the investment decision in the profit stage, we model this banking system as a two-stage problem, which is shown in Fig. 1.

In practice, some data could be imprecise. Total risk-weighted assets refer to the sum of In-Balance-Sheet assets and Off-Balance-Sheet exposures, weighted according to different risk coefficients, respectively. The risk coefficients are affected by macroeconomic environment, current risk coefficient, exposure of bank, etc., which is difficult to ascertain. The non-performing loans/assets are loans with signs of difficulties to repay the loans principal and interest. It is different to ascertain whether the borrowers can repay the loans on time, which are affected by public policy, credit of the clients, actual situation of the clients, etc. Due to the variability of risk coefficients and the repayment ability of borrowers, we employ fuzzy numbers to describe the total risk-weighted assets and non-performing loans/assets.

Definition 1. [21] Let L(·), R(·) be two reference functions. If the membership function of fuzzy variable ξ has the following form:

$$\mu_\xi(x) = \begin{cases} L(\frac{m-x}{\alpha}), & x \le m, \alpha > 0 \\ R(\frac{x-m}{\beta}), & x \ge m, \beta > 0. \end{cases}$$

Then ξ is called an LR fuzzy variable denoted by $(m, \alpha, \beta)_{LR}$, where m is the central value of ξ, α, β are the left and right spreads respectively of ξ. And the $\gamma - cut(0 \le \gamma \le 1)$ for the LR fuzzy variable is:

$$\xi_\gamma = (m - L^{-1}(\gamma)\,\alpha, m + R^{-1}(\gamma)\,\beta), \gamma \in [0, 1].$$

3 Models

Charnes-Cooper-Rhodes (CCR) model assumes all the DMUs are Constant Returns to Scale (CRS) [3]. In some cases we also have Variable Returns to Scale (VRS). In this section, three kinds of DEA fuzzy models are developed for commercial banking system evaluation.

3.1 Fuzzy CCR DEA Model with Undesirable Outputs

When the banking system is considered as a "black box", then we formulate the CCR model as follows.

$$
\begin{aligned}
&\max\ h_0^{\mathrm{CCR}} = \frac{u_d y_0^{RP}}{\sum_{i=1}^{4} v_i x_{i0} + w_u \tilde{z}_0^{\mathrm{RWA}} + u_u \tilde{y}_0^{\mathrm{NPA}}} \\
&s.t. \begin{cases} \dfrac{u_d y_j^{RP}}{\sum_{i=1}^{4} v_i x_{ij} + w_u \tilde{z}_j^{\mathrm{RWA}} + u_u \tilde{y}_j^{\mathrm{NPA}}} \leq 1, \quad j = 1, \cdots, J \\ v_i, w_u, u_d, u_u \geq \varepsilon > 0, \end{cases}
\end{aligned} \tag{1}
$$

where v_i, w_u, u_d, and u_u are non-negative weights. ε is a small non-Archimedean number imposed for avoiding ignorance of any factor. $x_{ij}(i = 1, 2, 3, 4)$ represent $x_j^{NE}, x_j^{FA}, x_j^{OE}, x_j^{NI}$, respectively.

3.2 Two-Stage Fuzzy CRS DEA Model with Undesirable Outputs

Next we discuss the case when the banking system is divided into two stages. Suppose all the DMUs are CRS, we develop models (2) and models (3) to calculate the efficiencies of the first stage and the second stage.

$$
\begin{aligned}
&\max h_0^{\mathrm{CRS}^1} = \frac{w_d z_0^{TD}}{\sum_{i=1}^{4} v_i x_{i0} + w_u \tilde{z}_0^{\mathrm{RWA}}} \\
&s.t. \begin{cases} \dfrac{w_d z_j^{TD}}{\sum_{i=1}^{4} v_i x_{ij} + w_u \tilde{z}_j^{\mathrm{RWA}}} \leq 1, j = 1, \cdots, J \\ v_i, w_d, w_u \geq \varepsilon > 0, \end{cases}
\end{aligned} \tag{2}
$$

and

$$
\begin{aligned}
&\max h_0^{\mathrm{CRS}^2} = \frac{u_d y_0^{RP}}{w_d z_0^{TD} + u_u \tilde{y}_0^{\mathrm{NPA}}} \\
&s.t. \begin{cases} \dfrac{u_d y_j^{RP}}{w_d z_j^{TD} + u_u \tilde{y}_j^{\mathrm{NPA}}} \leq 1, j = 1, \cdots, J \\ w_d, u_d, u_u \geq \varepsilon > 0. \end{cases}
\end{aligned} \tag{3}
$$

Then we use ω_1 and ω_2 to represent the weights of the first stage and the second stage, and the formulations are as follows.

$$
\omega_1 = \frac{\sum_{i=1}^{4} v_i x_{ij} + w_u \tilde{z}_j^{\mathrm{RWA}}}{\sum_{i=1}^{4} v_i x_{ij} + w_u \tilde{z}_j^{\mathrm{RWA}} + w_d z_j^{TD} + u_u \tilde{y}_j^{\mathrm{NPA}}}, \tag{4}
$$

and

$$
\omega_2 = \frac{w_d z_j^{TD} + u_u \tilde{y}_j^{\mathrm{NPA}}}{\sum_{i=1}^{4} v_i x_{ij} + w_u \tilde{z}_j^{\mathrm{RWA}} + w_d z_j^{TD} + u_u \tilde{y}_j^{\mathrm{NPA}}}. \tag{5}
$$

Then the two-stage fuzzy CRS DEA model with undesirable can be proposed as Eq. (6).

$$\max\ h_0^{\mathrm{CRS}} = \omega_1 h_0^{\mathrm{CRS}^1} + \omega_2 h_0^{\mathrm{CRS}^2} = \frac{w_d z_0^{TD} + u_d y_0^{RP}}{\sum_{i=1}^{4} v_i x_{i0} + w_u \tilde{z}_0^{\mathrm{RWA}} + w_d z_0^{TD} + u_u \tilde{y}_0^{\mathrm{NPA}}}$$
$$s.t. \begin{cases} \dfrac{w_d z_j^{TD}}{\sum_{i=1}^{4} v_i x_{ij} + w_u \tilde{z}_j^{\mathrm{RWA}}} \le 1, \quad j = 1, \cdots, J \\ \dfrac{u_d y_j^{RP}}{w_d z_j^{TD} + u_u \tilde{y}_j^{\mathrm{NPA}}} \le 1, \quad j = 1, \cdots, J \\ v_i, w_d, w_u, u_d, u_u \ge \varepsilon > 0. \end{cases} \tag{6}$$

3.3 Two-Stage Fuzzy VRS DEA Model with Undesirable Outputs

Often, we also face VRS. In order to consider the efficiency of the banks in the case of VRS, we propose the two-stage fuzzy VRS model as follows.

$$\max\ h_0^{VRS} = \frac{w_d z_0^{TD} + \mu_1 + u_d y_0^{RP} + \mu_2}{\sum_{i=1}^{4} v_i x_{i0} + w_u \tilde{z}_0^{RWA} + w_d z_0^{TD} + u_u \tilde{y}_0^{NPA}}$$
$$s.t. \begin{cases} \dfrac{w_d z_j^{TD} + \mu_1}{\sum_{i=1}^{4} v_i x_{ij} + w_u \tilde{z}_j^{RWA}} \le 1, \quad j = 1, \cdots, J \\ \dfrac{u_d y_j^{RP} + \mu_2}{w_d z_j^{TD} + u_u \tilde{y}_j^{NPA}} \le 1, \quad j = 1, \cdots, J \\ v_i, w_d, w_u, u_d, u_u \ge \varepsilon > 0 \end{cases} \tag{7}$$

If $\mu > 0$, then the DMU has increasing returns to scale; if $\mu = 0$, then the DMU has constant returns to scale; if $\mu < 0$, then the DMU has decreasing returns to scale.

4 Solution Method

In this section, we propose a solution to the two-stage DEA models with fuzzy parameters.

4.1 Solution Method Equivalent Two-Stage CCR DEA Model

First, applying the Charnes-Cooper transformation (C-C transformation), the two-stage model (8) can be converted into the following linear program, and the efficiency of DMU_0 can be calculated. The optimal efficiency g_j^{CCR} is equal to the optimal efficiency h_j^{CRS}.

$$\max\ g_0^{CCR} = u_d y_0^{RP}$$
$$s.t. \begin{cases} u_d y_j^{RP} - \sum_{i=1}^{4} v_i x_{ij} - w_u \tilde{z}_j^{RWA} - u_u \tilde{y}_j^{NPA} \le 0, \quad j = 1, \cdots, J \\ \sum_{i=1}^{4} v_i x_{i0} + w_u \tilde{z}_0^{RWA} + u_u \tilde{y}_0^{NPA} = 1 \\ v_i, w_u, u_d, u_u \ge \varepsilon > 0. \end{cases} \tag{8}$$

Because of the fuzzy parameters, we use chance constrained operator to handle model (8), and we get model. (9).

$$\max\ g_0^{\mathrm{CCR}} = u_d y_0^{RP}$$
$$s.t.\begin{cases} \mathrm{Pos}\left\{u_d y_j^{RP} - \sum_{i=1}^{4} v_i x_{ij} - w_u \tilde{z}_j^{\mathrm{RWA}} - u_u \tilde{y}_j^{\mathrm{NPA}} \le 0\right\} \ge \gamma, \quad j = 1, \cdots, J \\ \mathrm{Pos}\left\{\sum_{i=1}^{4} v_i x_{i0} + w_u \tilde{z}_0^{\mathrm{RWA}} + u_u \tilde{y}_0^{\mathrm{NPA}} \ge 1\right\} \ge \gamma \\ \mathrm{Pos}\left\{\sum_{i=1}^{4} v_i x_{i0} + w_u \tilde{z}_0^{\mathrm{RWA}} + u_u \tilde{y}_0^{\mathrm{NPA}} \le 1\right\} \ge \gamma \\ v_i, w_u, u_d, u_u \ge \varepsilon > 0 \end{cases} \tag{9}$$

Theorem 1. *[21] Given two LR fuzzy variables* $\xi_1 = (m_1, \alpha_1, \beta_1)$, $\xi_2 = (m_2, \alpha_2, \beta_2)$, *where* m *are the central values; and* α *and* β *are the left and the right spreads respectively. Then we have:*

$$k\xi_1 = (km_1, k\alpha_1, k\beta_1), k > 0$$
$$\begin{cases} \xi_1 + \xi_2 = (m_1 + m_2, \alpha_1 + \alpha_2, \beta_1 + \beta_2) \\ \xi_1 - \xi_2 = (m_1 - m_2, \alpha_1 + \beta_2, \beta_1 + \alpha_2) \end{cases}$$
$$Pos\{\xi_1 \ge \xi_2\} \ge \gamma \Leftrightarrow m_{1\gamma}^R \ge m_{2\gamma}^L,$$

where k *is a real number,* $\gamma(0 \le \gamma \le 1)$ *are the possibility values,* $m_{1\gamma}^R$ *is the right end point of* $\gamma - cut$ *of* ξ_1*, and* $m_{2\gamma}^L$ *is the left end point of* $\gamma - cut$ *of* ξ_2*.*

According to Theorem 1, model. (9) can be transformed into model. (10).

$$\max\ g_0^{CCR} = u_d y_0^{RP}$$
$$s.t.\begin{cases} u_d y_j^{RP} - \sum_{i=1}^{4} v_i x_{ij} - (w_u (z_j^{RWA})_m + u_u (y_j^{NPA})_m) - (1-\gamma)(w_u (z_j^{RWA})_\beta + u_u (y_j^{NPA})_\beta) \le 0, \\ j = 1, \cdots, J \\ \sum_{i=1}^{4} v_i x_{i0} + w_u (z_0^{RWA})_m + u_u (y_0^{NPA})_m + (1-\gamma)(w_u (z_0^{RWA})_\beta + u_u (y_0^{NPA})_\beta) \ge 1 \\ \sum_{i=1}^{4} v_i x_{i0} + w_u (z_0^{RWA})_m + u_u (y_0^{NPA})_m - (1-\gamma)(w_u (z_0^{RWA})_\alpha + u_u (y_0^{NPA})_\alpha) \le 1 \\ v_i, w_u, u_d, u_u \ge \varepsilon > 0 \end{cases} \tag{10}$$

After getting the optimal solutions $\mathrm{v}_r^*, \mathrm{w}_u^*, u_d^*$ *and* u_u^**, the CCR DEA efficiency* g_0^{CCR} *of the evaluated whole system can be obtained.*

4.2 Equivalent Two-Stage CRS and VRS DEA Models

Similarly, models (6) and (7) can be converted into models (11) and (12):

$$\max\ \mathrm{g}_0^{\mathrm{CRS}} = w_d z_0^{TD} + u_d y_0^{RP}$$
$$s.t. \begin{cases} w_d z_j^{TD} - \sum_{i=1}^{4} v_i x_{ij} - w_u (z_j^{\mathrm{RWA}})_m - (1-\gamma) w_u (z_j^{\mathrm{RWA}})_\beta \leq 0, j = 1, \cdots, J \\ u_d y_j^{RP} - w_d z_j^{TD} - u_u (y_j^{\mathrm{NPA}})_m - (1-\gamma) u_u (y_j^{\mathrm{NPA}})_\beta \leq 0, j = 1, \cdots, J \\ \sum_{i=1}^{4} v_i x_{i0} + w_d z_0^{TD} + w_u (z_0^{\mathrm{RWA}})_m + u_u (y_0^{\mathrm{NPA}})_m + (1-\gamma)(w_u (z_0^{\mathrm{RWA}})_\beta + u_u (y_0^{\mathrm{NPA}})_\beta) \geq 1 \\ \sum_{i=1}^{4} v_i x_{i0} + w_d z_0^{TD} + w_u (z_0^{\mathrm{RWA}})_m + u_u (y_0^{\mathrm{NPA}})_m - (1-\gamma)(w_u (z_0^{\mathrm{RWA}})_\alpha + u_u (y_0^{\mathrm{NPA}})_\alpha) \leq 1 \\ v_i, w_d, w_u, u_d, u_u \geq \varepsilon > 0, \end{cases} \tag{11}$$

$$\max\ \mathrm{g}_0^{VRS} = w_d z_0^{TD} + \mu_1 + u_d y_0^{RP} + \mu_2$$
$$s.t. \begin{cases} w_d z_j^{TD} - \sum_{i=1}^{4} v_i x_{ij} - w_u (z_j^{\mathrm{RWA}})_m - (1-\gamma) w_u (z_j^{\mathrm{RWA}})_\beta + \mu_1 \leq 0, j = 1, \cdots, J \\ u_d y_j^{RP} - w_d z_j^{TD} - u_u (y_j^{\mathrm{NPA}})_m - (1-\gamma) u_u (y_j^{\mathrm{NPA}})_\beta + \mu_2 \leq 0, j = 1, \cdots, J \\ \sum_{i=1}^{4} v_i x_{i0} + w_d z_0^{TD} + w_u (z_0^{\mathrm{RWA}})_m + u_u (y_0^{\mathrm{NPA}})_m + (1-\gamma)(w_u (z_0^{\mathrm{RWA}})_\beta + u_u (y_0^{\mathrm{NPA}})_\beta) \geq 1 \\ \sum_{i=1}^{4} v_i x_{i0} + w_d z_0^{TD} + w_u (z_0^{\mathrm{RWA}})_m + u_u (y_0^{\mathrm{NPA}})_m - (1-\gamma)(w_u (z_0^{\mathrm{RWA}})_\alpha + u_u (y_0^{\mathrm{NPA}})_\alpha) \leq 1 \\ v_i, w_d, w_u, u_d, u_u \geq \varepsilon > 0. \end{cases} \tag{12}$$

By solving models (11) and (12), the CRS and VRS DEA efficiencies of the evaluated whole system, $\mathrm{g}_0^{\mathrm{CRS}}$ and $\mathrm{g}_0^{\mathrm{VRS}}$, can be obtained. Also the efficiencies of each sub-stage $\mathrm{g}_0^{\mathrm{CRS}^1}$ ($\mathrm{g}_0^{\mathrm{VRS}^1}$) and $\mathrm{g}_0^{\mathrm{CRS}^2}$ ($\mathrm{g}_0^{\mathrm{VRS}^2}$) can be obtained.

5 Case Study

In this section, we evaluate the efficiencies of the top 16 Chinese banks: Bank Of China (BOC), Agricultural Bank of China (ABC), Industrial and Commercial Bank of China (ICBC), China Construction Bank (CCB), Bank of Communications (BCM), China Citic Bank (CNCB), Ping An Bank (PAB), Hua Xia Bank (HXB), China Everbright Bank (CEB), China Industrial Bank (CIB), China Merchants Bank (CMB), Bank Of Ningbo (NBB), Bank of Nanjing (NJB), Bank of Beijing (BJB), Shanghai Pudong Development Bank (SPDB), and China Minsheng Banking (CMSB).

5.1 Data

Table 1 presents the data of the inputs and outputs of the selected 16 Chinese commercial banks in 2015.

5.2 Results

Based on the data in Table 1, first we set the decision-makers' confidence levels for the two stages as $\gamma = 0.9$, then the efficiencies of each bank in different circumstances can be produced after solving two-stage fuzzy CCR, CRS, and VRS DEA models by LINGO, respectively. The results are shown in Tables 2 and 3. And Fig. 2 compares the system efficiencies of 16 commercial banks based on the proposed three kinds of models.

Table 1. Data

Bank	Number of employees (×100)	Fixed assets (100 million RMB)	Operating expenses (100 million RMB)	Number of institutions (×10)	Total deposits (10 billion RMB)	Total risk-weighted assets (10 billion RMB)	Retained profits (100 million RMB)	Non-Performing loans/Assets (100 million RMB)
BOC	3100	1820	2439	1069	1173	(1065,51,38)	1794	(1,309,181,120)
ABC	5031	1562	3033	2367	1354	(1099,13,54)	1808	(2,129,706,125)
ICBC	4663	2215	1024	1673	1628	(1322,55,88)	2777	(1,795,355,168)
CCB	3692	1595	3091	1492	1367	(1072,37,52)	2289	(1,660,372,160)
BCM	915	904	1081	314	448	(465,33,29)	668	(562,96,52)
CNCB	499	160	905	135	318	(347,41,27)	417	(361,66,25)
PAB	323	479	673	100	173	(166,18,25)	219	(176,45,41)
HXB	342	113	339	79	135	(133,9,14)	190	(163,43,24)
CEB	403	126	539	94	199	(219,23,30)	296	(244,65,11)
CIB	520	178	915	179	248	(343,37,16)	507	(260,61,67)
CMB	762	308	1272	171	357	(321,22,0)	580	(474,136,78)
NBB	95	502	115	29	36	(44,7,7)	66	(24,4,3)
NJB	74	450	139	15	50	(50,8,10)	71	(21,3,6)
BJB	138	713	230	44	102	(128,15,9)	169	(87,24,9)
SPDB	484	1906	4629	166	295	(337,36,24)	510	(351,104,73)
CMSB	595	3773	942	106	273	(335,35,13)	470	(328,90,53)

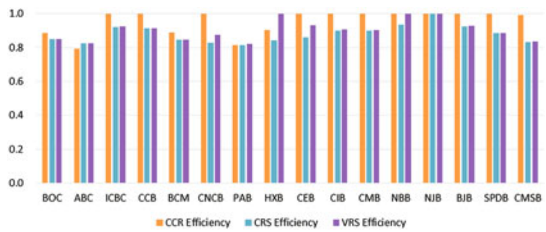


Fig. 2. Efficiencies based on CCR, CRS and VRS models

5.3 Comparison and Discussion

According to Table 2, ICBC, CCB, CNCB, CEB, CIB, CMB, NBB, BJB, SPDB, and CMSB are at the frontier when using traditional CCR DEA model evaluation, but the advantages are not maintained when using CRS DEA model. At the same time, the efficiencies of BOC, BCM, HXB, CMSB also decline in varying amounts. We conclude that often the CCR model may get high efficiency scores when the system is considered as a "black box" by using four inputs (Employees, Fixed assets, Operating expenses and Institutions) to produce three outputs (Total Risk-Weighted Assets, Retained profits and Non-Performing loans/Assets). However, the efficiencies may be not that high any more when the intermediates are taken into account. In other words, the traditional CCR model can only find weak effective system, but it is different to distinguish which banks are truly more efficient.

Table 2. Efficiencies based on two-stage CCR and CRS models

Bank	CCR DEA efficiency	CRS two-stage DEA efficiency			Bank	CCR DEA efficiency	CRS two-stage DEA efficiency		
		System	Stage 1	Stage 2			System	Stage 1	Stage 2
BOC	0.89	0.85	0.95	0.75	CEB	1	0.86	0.99	0.73
ABC	0.79	0.82	0.99	0.66	CIB	1	0.9	0.78	1
ICBC	1	0.92	1	0.84	CMB	1	0.9	1	0.8
CCB	1	0.91	1	0.82	NBB	1	0.94	0.75	1
BCM	0.89	0.85	0.96	0.73	NJB	1	1	1	1
CNCB	1	0.83	1	0.64	BJB	1	0.93	1	0.86
PAB	0.81	0.81	1	0.62	SPDB	1	0.89	0.92	0.85
HXB	0.9	0.84	1	0.69	CMSB	0.99	0.83	0.82	0.84

Table 3. Efficiencies under 2 return scale assumptions

Bank	CRS two-stage DEA efficiency	VRS two-stage DEA efficiency			Bank	CRS two-stage DEA efficiency	VRS two-stage DEA efficiency		
		System	Stage 1	Stage 2			System	Stage 1	Stage 2
BOC	0.85	0.85	0.95	0.75	CEB	0.86	0.93	1	0.74
ABC	0.82	0.82	0.99	0.66	CIB	0.9	0.91	0.8	1
ICBC	0.923	0.924	1	0.84	CMB	0.9	0.91	1	0.8
CCB	0.914	0.915	1	0.82	NBB	0.94	1	1	1
BCM	0.85	0.85	0.96	0.73	NJB	1	1	1	1
CNCB	0.83	0.88	1	0.65	BJB	0.926	0.93	1	0.86
PAB	0.81	0.82	0.97	0.63	SPDB	0.89	0.89	0.92	0.85
HXB	0.84	1	1	1	CMSB	0.834	0.835	0.82	0.85

By using two stage DEA model, we can further find out more discriminating DEA efficiencies for DMUs because of the stronger restriction on the production possible set. Although ICBC, CCB, CNCB, PAB, HXB, CMB and BJB are DEA efficient at the production stage, all the efficiencies of them at the profit stage are below 1. Similarly, CIB and NBB are inefficient at the profit stage while both of them are DEA efficient at the production stage. That is due to CCR DEA model neglects the internal interaction between the two stages. It also proves that model (6) reflects not only the efficiency of each stage, but also the efficiency of the banking system considering the relationship of two sub-stages.

Interestingly, the efficiency of ABC is improved by using model (6), which manifests ABC have a great advantage of the internal operational process. We find that the ratio of risk-weighted asset and inputs of the NJB is relatively small, the ratio of deposit and input is relatively high, and the input/output ratio of the NJB is at a reasonable level. Further analysis shows that both efficiency values of NJB in model (1) and model (6) are 1, i.e. they are both located in the frontier. This is consistent with our analysis, and policy makers can consider them as the industry benchmark.

If all the inputs of the bank have changed in the same proportion, this change will have an impact on the total outputs of the system. That is, the banking system will have VRS in the actual operation process. Based on above, we propose two kinds of models under the CRS and VRS assumptions. According to the model (7), the corresponding DEA efficiency can be obtained in Table 3. In Table 3, column 2 and column 3 are the two-stage banking system overall efficiency values obtained under the two assumptions, respectively.

It is found that HXB, NBB and NJB are DEA efficient not only in the whole bank system, but also in both of the two sub-stages. Compared with two-stage CRS DEA model, the efficiencies of two-stage VRS DEA increase slightly, and the number of points on the frontier also increases. Accordingly, we can draw a conclusion that commercial banks will have better performance in the VRS conditions. Therefore, in order to obtain a reasonable input/output combination to improve the performance of the system, the decision-makers can improve the scale of return by some actions (such as training staff to improve business efficiency, upgrading of banking equipment, improve its business throughput, and so on). In general, the difference between DEA efficiencies under the assumption of different scale returns is in a small range, which proves that the models we proposed have certain stability.

6 Conclusion

This paper explores a problem of efficiency evaluation of banking system under an uncertain environment. We consider the bank operating process consists of production stage and profit stage, and the two sub-stages of system are connected with each other by intermediate factor. Funds are collected from depositors in the production stage, which are used to invest into other activities to get profits in the profit stage. The undesirable outputs of both two sub-stages are treated as inputs. On this basis, we present two-stage fuzzy DEA models with undesirable outputs based on two different returns to scale.

In this paper, we have three main contributions as follows: First, based on two assumptions of returns to scale, we present the two-stage fuzzy CRS and VRS DEA models with undesirable output. The efficiency of the whole system is defined as the weighted sum of the efficiencies of two sub-stages, where the weights depend on the importance of the two stages. Second, considering the variability of risk coefficients and the repayment ability of borrowers, we apply fuzzy numbers to describe the total risk-weighted assets and non-performing loans/assets, which is closer to practice. Then we apply chance constrained operator to handle the fuzzy factors, and the objective is optimized subject to the chance constraints under certain confidence levels. Third, a case study and detailed comparison discussion are given by examining the top 16 China's commercial banks in 2015.

This paper focuses on the actual operation of commercial banks, considering the internal operational structure of commercial banks and its undesirable output in the process of operation. Through the application of the model, it is found that the model can better evaluate the input/output structure of the bank and has the ability of pointing out the direction to be improved.

Acknowledgements. We thank those that have given constructive comments and feedback to help improve this paper. Support was provided by the National Natural Science Foundation of China (Grant No. 71401093), China Scholarship Council (Grant No. 201606875006), Soft Science Research Project of Shaanxi Province (Grant No. 2016KRM089), Research Center for Systems Science & Enterprise Development (Grant No. Xq16B01), and the Fundamental Research Funds for the Central Universities (Grant Nos. 14SZYB08, WUT: 2016VI002).

References

1. Akther S, Fukuyama H, Weber WL (2013) Estimating two-stage network slacks-based inefficiency: an application to bangladesh banking. Omega 41(1):88–96
2. An Q, Chen H et al (2015) Measuring slacks-based efficiency for commercial banks in China by using a two-stage DEA model with undesirable output. Ann Oper Res 235(1):13–35
3. Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444
4. Chen Y, Cook WD, Zhu J (2010) Deriving the DEA frontier for two-stage processes. Eur J Oper Res 202(1):138–142
5. Cook WD, Liang L, Zhu J (2010) Measuring performance of two-stage network structures by DEA: a review and future perspective. Omega 38(6):423–430
6. Hatami-Marbini A, Saati S, Tavana M (2010) An ideal-seeking fuzzy data envelopment analysis framework. Appl Soft Comput 10(4):1062–1070
7. Hatami-Marbini A, Emrouznejad A, Tavana M (2011) A taxonomy and review of the fuzzy data envelopment analysis literature: two decades in the making. Eur J Oper Res 214(3):457–472
8. Kao C, Liu ST (2011) Efficiencies of two-stage systems with fuzzy data. Fuzzy Sets Syst 176(1):20–35
9. Khodabakhshi M, Gholami Y, Kheirollahi H (2010) An additive model approach for estimating returns to scale in imprecise data envelopment analysis. Appl Math Model 34(5):1247–1257
10. Lertworasirikul S, Fang SC et al (2003) Fuzzy data envelopment analysis (DEA): a possibility approach. Fuzzy Sets Syst 139(2):379–394
11. Liu ST (2014) Restricting weight flexibility in fuzzy two-stage DEA. Comput Ind Eng 74:149–160
12. Liu W, Zhou Z et al (2015) Two-stage DEA models with undesirable input-intermediate-outputs. Omega 56:74–87
13. Puri J, Yadav SP (2013) A concept of fuzzy input mix-efficiency in fuzzy DEA and its application in banking sector. Expert Syst Appl 40(5):1437–1450
14. Puri J, Yadav SP (2014) A fuzzy DEA model with undesirable fuzzy outputs and its application to the banking sector in India. Expert Syst Appl 41(14):6419–6432
15. Scheel H (2001) Undesirable outputs in efficiency valuations. Eur J Oper Res 132(2):400–410
16. Sengupta JK (1992) A fuzzy systems approach in data envelopment analysis. Comput Math Appl 24(8):259–266
17. Tavana M, Khalili-Damghani K (2014) A new two-stage stackelberg fuzzy data envelopment analysis model. Measurement 53:277–296
18. Wanke P, Barros C, Emrouznejad A (2016) Assessing productive efficiency of banks using integrated fuzzy-DEA and bootstrapping: a case of mozambican banks. Eur J Oper Res 249(1):378–389

19. Wu DD, Yang Z, Liang L (2006) Efficiency analysis of cross-region bank branches using fuzzy data envelopment analysis. Appl Math Comput 181(1):271–281
20. Hu XY, Cheng XJ, Ma LJ (2013) Efficiency evaluation of commercial banks based on two-stage DEA model considering undesirable outputs. J Univ Chin Acad Sci 30(4):462–471 (in Chinese)
21. Xu J, Zhou X (2011) Fuzzy-like multiple objective decision making, vol 263. Springer

Increasing Effect in Lodger Number of Hot Spring Hotel According to the Started Operation of Hokuriku Shinkansen

Takashi Oyabu[1](✉) and Junko Nakamura[2]

[1] Kokusai Business Gakuin College, Kanazawa, Japan
oyabu24@gmail.com
[2] Kanazawa Seiryo University, Kanazawa, Japan

Abstract. In Hokuriku district, economic efficiency is improved according to the stated operation of Hokuriku Shinkansen. However, there are some variations in the effect each area. It is necessary to utilize the effect into a future tourism strategy by indicating the effects. In this experiment, the effect is evaluated using the characteristics of the lodger number in nine spa areas of Hokuriku district. As a result, there is a larger effect as an area is near to a Shinkansen station and the increasing rate is over 30%. Other areas except Wajima area have the rates of 15 to 20%. The one for Wajima area is 33%. It keeps at a distance from Kanazawa which is the nearby station of Wajima area. Popular NHK television-drama (title is "Mare") was televised for end of March to September 2015 (half-year). Wajima area was the filming site. It is estimated that there is an effect over 10% as the filmed site. There is a large effect by connecting Hokuriku with Tokyo area by Shinkansen (high speed railway).

Keywords: Hokuriku Shinkansen · Hot spring · Lodger number · Economic effect

1 Introduction

Japan's population has entered into a decreasing phase due to the decreasing birthrate and aging population. Productive population also decreases remarkably, therefore a measure to put on the brake for the phenomenon is required. National budget is also tight due to the increasing of the cost of social security including medical expense as the society ages [2]. It is essential to build a policy which could create a new industry and put the brakes on the depopulation in a rural area. One of the regional-revitalization measures is "tourism industry". Human exchange is a basic policy to develop the tourism industry. The following measures are important in the exchange, namely information and transportation infrastructures [4].

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_136

Comfortable and express railway system (e.g. Shinkansen) had been desired in Hokuriku district (Fukui, Ishikawa and Toyama Prefectures) for 50 years. The newest section of the Hokuriku Shinkansen Line, between Nagano and Kanazawa, opened on March 14, 2015. The line was opened about fifty years after Tokaido Shinkansen Line started operations. The number of visitors was increased considerably due to the operation in Kanazawa City (Shinkansen effect), which was the prefectural capital of Ishikawa. The occupancy rate for hotels was over 85%. There are nine main spa areas mainly in Hokuriku district, namely Awara, Yamanaka, Yamashiro, Katayamatsu, Awazu, Yuwaku, Wakura, Wajima and Unazuki. The number of lodgers has increased nearly 15 to 30% in the areas. The areas except for Unazuki (Toyama Prefecture) and Yuwaku (Kanazawa City) take one hour and half by bus from each Shinkansen station. Unazuki and Yuwaku take within thirty minutes by car or train. The nearby station of Unazuki is Kurobe-Unazuki-Onsen and the one for Yuwaku is Kanazawa. The number of visitors for both spa areas increased about 30% and the start of the operation significantly affected the number for both areas. It is desirable for the ripple effect to affect whole areas. The mass media reported that the number increased in whole spa areas. It is necessary to investigate the effect and apply the survey result to the future strategy.

Outline of ripple effect can be known by examining the monthly variability of lodger number in each spa area and it is possible to determine the date of various kinds of providing events [1,3,11]. Moreover, it is helpful to construct the effective strategy that increases the guest number. The variability of lodger number in nine spa areas of Hokuriku District for three years is summarized as Shinkansen effect in this study. The data for December 2015 are estimated by the authors.

2 Hokuriku Shinkansen

Hokuriku Shinkansen is a railway rout to connect Tokyo and Osaka via Hokuriku District at the entire line available. When the entire rail line of Hokuriku Shinkansen is opened, it can assume the bypath function of Tokaido Shinkansen. A part of the line (from Tokyo to Nagano) opened in ahead of schedule in 1997 due to Nagano Olympics (1998 Winter Olympics). It took 18 years until Kanazawa opening of business. It was fifty years late than the opening of Tokaido Shinkansen (opened in 1964, Tokyo to Shin-Osaka). The line was the world's first high speed railway. The operation of Shinkansen line connected to Tokyo was desired earnestly for the tourism promotion and economic development in the district. The passenger number increased triple than the previous one. However, airplane passenger decreased sharply (more than 30%) and the number of the flights (including number of seats) was decreasing after the start of the Hokuriku Shinkansen Line. There are three airports in Hokuriku, namely Komatsu (Ishikawa Pref.), Noto (Ishikawa) and Toyama (Toyama). The maximum speed of the line is restricted to 260 km/hour for noise reduction and it is investigated to speed up (300 km/hour) in consideration of noise environment.

The maximum speed of each Shinkansen Line is indicated in Table 1. The speed of Tohoku Shinkansen is the fastest (320 km/hour). Speedup is important in future. Photograph of Hokuriku Shinkansen is shown in Fig. 1. Sky blue and copper color of the body provide a calm atmosphere.

Table 1. Maximum speed for each Shinkansen in Japan

Shinkansen	km/hour
Tokaido	285
Sanyo	300
Tohoku	320
Jyoetsu	240
Kyushu	260
Hokuriku	260
Hokaido	260

Fig. 1. Hokuriku Shinkansen

The route of Hokuriku Shinkansen is indicated in Fig. 2. The line will be expanded to Shin-Osaka and connected to Tokaido Shinkansen via Koto.

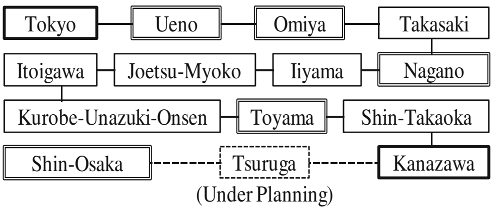


Fig. 2. The route of Hokuriku Shinkansen

It will take about twenty years [5]. The Linear Chuo Shinkansen will be able to run between Tokyo and Nagoya in just 40 min in 2027. The line will be expanded to Osaka from Nagoya in 2045 and it will take 67 min between Tokyo and Osaka. Strategy to take the interchange of the person is necessary, but it is also necessary to remark the comfortability, safety and security besides speed. Japan developed firstly a rapid transit railway in the world. Moreover, noise and environmental assessment should be remarked. Hokuriku Shinkansen is a pivot of the regional revitalization.

About 8.15 million passengers used Hokuriku Shinkansen in total since started operation (for Mar. 14, 2015 to Jan. 20, 2016; about 10 months). The number was about 3 times than heretofore and the handling incomes increased by 33%, which was released by JR West. The change of passengers for Mar. 14 to Sep. 13 (six months) is indicated in Fig. 3. The data was measured for Jyoetsu-Myoko to Itoigawa (working area of JR west). The values in March and September are very small because they are in half month. The average for April to August is about 0.8 million and the number of passengers a day each station is represented in Table 2. The one for Kanazawa is less than twice the one for Toyama.

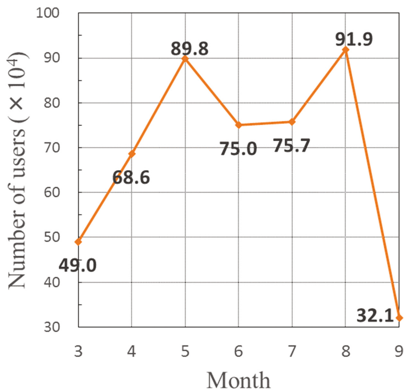


Fig. 3. Number of passenger of Hokuriku Shinkansen (Jyoetsu-Myoko to Itoigawa), which is measured for Mar. 14 to Sep. 13, 2015)

Table 2. Number of passengers a day in five stations

Station	Baoding number/day
Kanazama	8930
Toyama	4860
Shin-Takaoka	1600
Unazuki	840
Itoigawa	400

3 Number of Guest in Hokuriku Spa Areas

Hot spring is an important tourism resource and the economic effect by the guests is especially large in Hokuriku District. There are nine spa areas as main hot spring resorts in the district (Fukui, Ishikawa and Toyama Prefectures). In this study, Shinkansen effect was evaluated using the number of hotel guest in the areas. The number is introduced by the guests of the facilities which are joining Ryokan association (inn association). Only a few facilities do not join the association and the unjointed number is less than 10%. The data is not correspond to the local overall hotel guests. The nine survey target areas are as follows:

Awara (Awara, Fukui Pref.), Awazu (Komatsu, Ishikawa Pref.),
Yamanaka, Yamashiro and Katayamazu (Kaga, Ishikawa Pref.)
Wajima (Wajima, Ishikawa Pref.), Wakura (Nanao, Ishikawa Pref.)
Yuwaku (Kanazawa, Ishikawa Pref.), Unazuki (Kurobe, Toyama Pref.)

Yamanaka, Yamashiro and Katayamazu are collectively called Kaga Onsen-Kyo (Kaga three major spas). Unazuki (15 km) and Yuwaku (20 km) are close to a Shinkansen station and it takes less than half hour from each station to these spa areas by bus or car. It takes over one and half hours form each near station for other areas by train or bus.

3.1 Kaga Four Major Spa Areas

Kaga three major spas (Kaga Onsen-Kyo) and Awazu spa are collectively called 'Kaga Four Onsen'. Awazu is located in Komatsu city which is neighbor of Kaga city. There are many visitors from Kansai area in the four spa areas. Osaka is the center of Kansai area in which the area population is about 20 million. There are also many visitors from Chukyo area (Nagoya) in which the area population is about 15 million. The number of visitor is decreasing gradually due to the slump of the Japanese economy in the areas and some hot-spring hotels discontinue their business. A cheap hotel chain is also expanding. Kaga three major spas (Kaga Onsen-Kyo; Yamanaka, Yamashiro and Katayamazu) had over 4 million visitors in about 1990. It was the maximum value. The number increased a little in 2015 than the one in 2014. It had about 2 million in 2015 due to the Shinkansen effect. Awazu had 0.17 million visitors in 2015 and the number was about 10% of the total of Kaga three major spas. These four spa areas (Kaga four major spa areas) are developing a tourism strategy comprehensively. The capacity of the four areas is about 15 thousands and there is a loop bus (named CANBUS) which connects the four areas. The area is unlimited ride by showing the ticket for the tourist. The ratio of departure area of the visitor to the areas was as follows before the opening of the Hokuriku Shinkansen Line (in February 2015), namely Kansai area is 32% and Kanto area is 4%. The one for Kanto approximately quadrupled and the one for Kansai was 23.5% in June 2015. There was the Shinkansen effects obviously [8]. The effect is high in Yamashiro spa area especially.

The characteristics of lodger number for three years (2013–2015) in the area are shown in Fig. 4. The values in 2013 and 2014 are 1.73 and 1.74 million. They are nearly equal. The percentages in the areas are as follows, Yamanaka: 25, Yamashiro: 39, Katayamazu: 27 and Awazu: 9%. The value for Yamashiro is maximum. The mean total visitors for April to November in 2013 and 2014 were derived and compared with the one for the same term in 2015 after the Shinkansen opening of business. It increased by 16% after the opening. The percentage is thought as the Shinkansen effect. The total lodger number for 2015 in the area is about 2 million. The characteristic of the lodger in number in the spa areas of Hokuriku District has a peak in August. The tendency is the same as the characteristic in Japan. However, there are also small peaks in March and November and the characteristics resemble each other depending on tourism resources. The visitors from Kanto area (area population: 42 million) increased to double after the start of the Hokuriku Shinkansen Line but the tendency was the same.

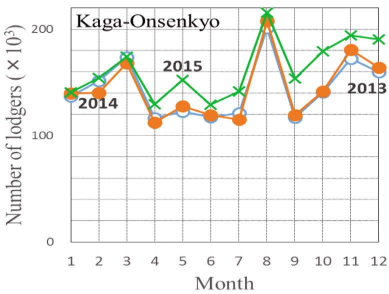


Fig. 4. Characteristic of lodger number for Kaga four major spa areas

3.2 Awara Spa Area

Awara spa area is located in Awara City (population: about 30 thousand) of Fukui prefecture and is called as a spa-resort of Kansai region. The city adjoins Kaga City. Therefore, a unified tourism strategy is sometimes built as the area of Kaga Onsen-Kyo altogether (including Yamashiro, Ymanaka, Katayamazu, Awazu and Awara). There are few tourism resources except the hot spring in Awara. The characteristics of lodger number for three years (2013–2015) in Awara spa area is shown in Fig. 5. The increasing percentage for April to November 2015 was 18%. The same deriving process was mentioned in the last section. The percentage is larger than the one for Kaga four major spa areas. The numbers in 2013 and 2014 were 0.61 and 0.69 million. It took a peak value (87 thousand) in August. The characteristics for Figs. 4 and 5 closely resemble to each other and it is a feature in Kaga Onsen-Kyo. The correlation diagram for both figures is shown in Fig. 6. The correlation coefficient R is 0.91 and the data

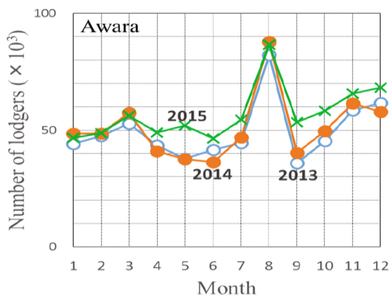


Fig. 5. Lodger characteristics in Awara spa area for 2013–2015

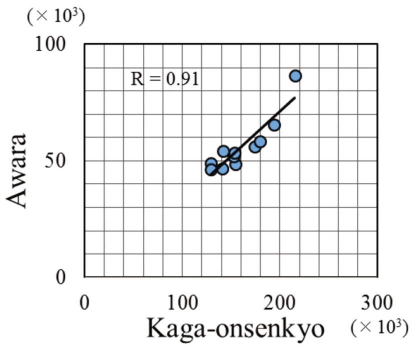


Fig. 6. Correlation diagram of lodger number between Kaga four major spa areas and Awara spa area

has a very high correlational relationship. If the coefficient is small, it can supplement the seasonal food and hotel guest mutually in the region. It is desirable for each facilities to have a specific tourism resources. And it is desirable that R is small. Lodger characteristics in five major spa areas of Kaga Onsen-Kyo should be varied due to a proposed strategy.

3.3 Yuwaku Spa Area

Yuwaku spa is located in Kanazawa city (population; 0.46 million) and is called as an inner room of Kanazawa. Business firms often give parties there to entertain their customers. There is Yumeji Takehisa Kan (a kind of museum) and many Yumeji-fans visit there. There is Lake Gyokusenko which is suited in the inner part of the hot-spring street. There is Himuro House (icehouse) by the lake and the winter snow is stored in the house. The stored operation is carried out in June every year. The snow ice was offered to Tokugawa family by Kaga Domain in the Edo Period. The many citizens have the habit of eating Himuro manju (steamed bun) to pray for good health at July 1. The area was the location of Hanasaku-iroha (an anime drama) broadcast on TV in 2011 and there were many visitors as the place of pilgrimage. Characteristic of lodger number in Yuwaku spa area is indicated in Fig. 7. The fluctuation for Yuwaku area is smaller than the ones for Kaga Onsen-Kyo spa areas which have a marked peak in August. The Shinkansen effect of 38% is derived, which is the biggest growth rate in this investigation. It is thought that Yuwaku is relatively near Kanazawa station as the reason for the high growth rate. The guests drift to the area because the occupancy rate for hotels near Kanazawa station is sometimes over 90%.

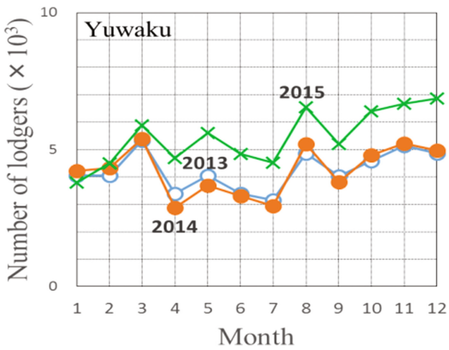


Fig. 7. Characteristic of lodger number in Yuwaku spa area

3.4 Wakura Spa Area

Wakura spa area is an entrance of the Noto sightseeing. It was said that the opening of the hot spring or bath house was about 1,200 years ago. A fisherman found a damaged white heron healing the wound in the hot spring. It was the origin. And there is a traditional Japanese inn "Kagaya" in the area. Kagaya was receiving repeatedly the best award as the most evaluated inn in Japan for 36 years until 2016. The award "100 selections of the best Japanese hotels and inns" is recognized by Ryoko-shinbun (a trip newspaper publisher). It is said that the hospitality (Omotenashi in Japanese) of Kagaya is number one in Japan, and the inn is also a place of corporate training by many companies. A picture of sending guest off is shown in Fig. 8. Wakura spa area is developing centered on Kagaya. Characteristic of lodger-number transition is presented in Fig. 9. It has the same trend as the one for Kaga Onsen Kyo, which is indicated in the figure and the increase of 20% is derived as Shinkansen effect. The decrease in number of the guests is relatively small in winter.

3.5 Wajima Spa Area

There are some Japanese inns and Minshukus (private home that runs inn providing room and board) in Wajima area but the capacity of guests is small. The number of guests was about 0.13 million in a year before the Hokuriku

Fig. 8. Hospitality of Kagaya (send-off)

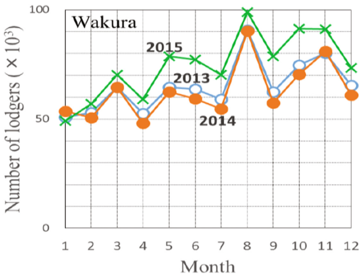


Fig. 9. Lodger number as a function of month in Wakura spa area

Shinkansen Line started operation. It became 0.17 million in 2015 and the increase of 30% is recognized. The increase rate of lodger number for April to November in 2015 was 33%. The number of visitors for Wajima morning market increased by 30%. The influence of the serial TV drama "Mare" produced by NHK is also thought. Wajima was used as a location for making the drama. It was broadcasted for March to September in 2015. There were really many visitors to Osawa area in Wajima City. Magaki (board fence) which is made up of bamboos, is famous in the area. Magaki protects a house from a strong winter wind (sea breeze). Lodger-number characteristic for Wajima spa area is presented in Fig. 10. The number increased after the broadcasted TV drama (May). The details for the increased number (30%) are as follows. Shinkansen effect for Wakura spa nearby Wajima is 20%. So Shinkansen effect is estimated by 20%, TV drama effect is 10%. It is difficult to judge the effect clearly because there is a synergistic effect.

3.6 Unazuku Spa Area

Unazuki is located to the east of Toyama prefecture and is the biggest hot spring resort in the prefecture. The name of the closest Shinkansen station to Unazuki is 'Kurobe-Unazuki Onsen'. There are a private railway from the station to Unazuki spa resort and it takes about 20 min by the train. There are a truck

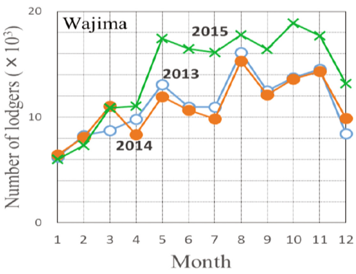


Fig. 10. Lodger-number characteristic for Wajima area

train in Unazuki to visit Kurobekeikoku (canyon). Hundreds of thousands of people visit the canyon annually. The scenery is wonderful especially for the autumn leaves. As for the result, the increasing rate of lodger number for April to November (after the Shinkansen opening of business) is over 30% and the shortage of human resources for some inns are remarkable in the area due to the increase of the room occupancy rate. The transition of lodger number is presented in Fig. 11. The increase after April is remarkable because the area is near from the Shinkansen station as well as Yuwaku. It is necessary to hold an attractive event in June and equalize the lodger number over whole year.

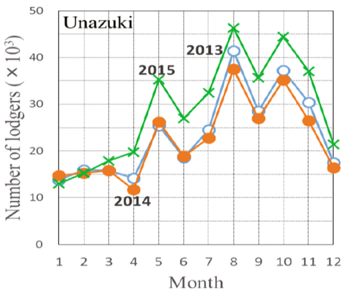


Fig. 11. Lodger number for Unazuki spa resort

3.7 Trend of Lodger Number Due to Shinkansen Started Operation

There were various kinds of advantages and disadvantages due to the Hokuriku Shinkansen started operation. The start of the operation spreads to over the entire area of Hokuriku (three prefectures) but there are the light and shade of the effect [8]. The movement of person is sucked into an area where an attractive event is held. This phenomenon is called "straw effect". Some points to be improved were listed by visitors to the area. The points are as follows.

- It is hard to make the reservation of the hotel.
- Hospitality is light due to the labor shortage.
- Rising of the meal price.

The increasing rates of lodger number for six hot spring areas as described above, is indicated altogether in Table 3. The rates of three areas (Yuwaku, Wajima and Unazuki) are over 30%. Yuwaku and Unazuki are near the Shinkansen stations, and Wajima is the filmed place for the drama (Mare). Consequently, the one for Wajima will decrease gradually. The correlation diagram for the lodger number between Yuwaku and Unazuki is derived and indicated in Fig. 12. The correlation coefficient R is 0.75 and the value is high. It is thought that there are some different types of tourism resources for the visitors in the area, for example autumn color of canyon, ski slope and cheap accommodation fee.

It was thought that the visitor number to Wajima increased by 10% due to the place for the TV drama "Mare". There was a synergistic effect with Shinkansen began the operation (increased by about 20%) and many persons visited Wajima to see the scene sites of the drama. Tokyo where the area population is 42 million, is the starting station of the Hokuriku Shinkansen. There is high potential to increase the number when the access becomes easy. The increasing rates of the five spa resorts (Kaga three major spa areas, Awazu and Awara) were almost 20%. The one for Wakura was 21%. The access for Wakura is relatively good because Wakura is connected directly to Kanazawa by JR railway and there is also a bus route. The rate is higher than the ones for the above mentioned five areas.

Coefficients of variation (cv: standard deviation/the mean value) for six spa resorts are summarized in Table 4. The values in 2014 and 2015 are represented. The ones for Wajima and Unazuki are exceed 0.3, namely the fluctuations for the lodger number of every month are larger than the other areas. A variation of the occupancy rate for each hotel is large, namely maintenance cost for the facilities including labor cost becomes higher. The coefficient for Wajima becomes large.

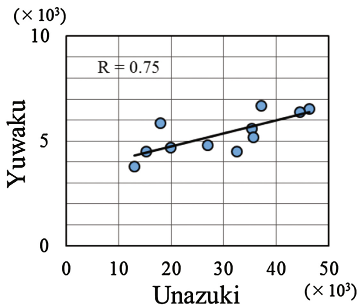


Fig. 12. Correlation diagram for the lodger number between Yuwaku and Unazuki

Table 3. Increasing rates for six hot spring areas in Hokuriku district

Hot spring	Kaga	Awara	Yuwaku	Wakura	Wajima	Unazuki
Rate	0.16	0.18	0.38	0.21	0.33	0.30

Table 4. Coefficients of variation for six spa resorts in 2014 and 2015

	Kaga	Awara	Yuwaku	Wakura	Wajima	Unazuki
2014	0.20	0.26	0.20	0.19	0.23	0.36
2015	0.16	0.20	0.17	0.20	0.31	0.38

4 Visitor Number for Other Main Facilities

There was a tourism spot in which the visitor number increased remarkably. It is Asakura remains in Ichijodani and located in Fukui prefecture. A CM movie was filmed in the spot and about 1.1 million persons visited the spot [7]. It is necessary to construct some facilities for visitors, namely parking place, rest house and souvenir shop. The elderly interested in history are also visiting. The service introducing the situation in ancient or olden times using a tablet is carried out.

Kenrokuen (one of the three largest gardens in Japan) is the spot which has the most visitors in Ishikawa prefecture and the number is 2.9 million in 2015 (increased by 45.5% than the one in 2014). There was 2.25 million visitors (increased by 81% compared with the previous year) in Kanazawa Castle Park which is adjacent to Kenrokuen. The number of foreign people visited Kenrokuen was 0.292 million (increased by 24.4% compared with the previous year). The visitors from Taiwan occupied by 46.5%. The one from Hong Kong was 10% and the one from USA was 5.3% (China was 5%). The rate of the one from Korea was few [6].

There is the Noto Railway Corporation in Noto area where Wajima and Wakura (in Nanao city) are located. It is a third sector railway company. The service is operating between Wakura and Anamizu (about 40 min by train), and JR operates a train between Nanao and Wakura. Many group tourists use the Noto Railway Corporation after the Hokuriku Shinkansen line started its operation. Some events were hold in the train, for example famous sweets-service and special guide by a conductor in the train. The events and TV drama "Mare" affected to the increase of the visitor number. The number was 61 thousand in 2015 (increased by 73% compared with the previous year). The visitor number form Kanto area were 42 thousand (70%). It is very large percentage compared with the ones for other areas [10]. The number for Tateyama Kurobe Alpine Route was 0.997 million in 2015 (increased by 10% compared with the previous year). There was a big snow wall (named Yuki-no-Otani in Japanese) in the route. One million visitors were expected as a goal in 2015. It was slightly few. As the details, the number of Japanese was 0.782 million and the one for foreign visitors was 0.215 million (increased by 12%). The number of visitors from Taiwan

(132 thousand) is the biggest and the second is Hong Kong (25 thousand) followed with Thailand (17 thousand) and Korea (16 thousand). Higher two countries are the same as Kenrokuen [9].

5 Conclusion

It is necessary to examine the visitor number after the Hokuriku Shinkansen line started its operation and to apply the result to a future strategy in the district. It is said that the number increases remarkably in the first year after the opening. However, it is also said that the effect does not continue after the second year. This phenomenon is called “Shinkansen jinx”. Some proper and attractive events are required to break the jinx every season and to secure the repeaters. Peculiar local hospitality (“Omotenashi” in Japanese) is also necessary to secure the repeaters. Some travel agents say that the traffic infrastructure and the accommodation cost (lodging and transportation expenses) are more important for repeaters. Some regional inhabitants have the objection against the increasing the number of visitors sharply because the community is in disorder. It is necessary to constitute the area in which the habitants and visitors can live together. The action of “Civic Pride” is also effective.

In this study, the fluctuation of the visitor number in spa resort areas is examined as a Shinkansen effect. It is important to understand the fluctuation and is also necessary for forming the social environment of the area in future. The visitor number of nine spa areas in Hokuriku district were investigated. As a result, it becomes obvious that the effect of visitors spreads to Hokuriku district whole. Especially, the lodger number of the areas near the two Shinkansen stations increased by about 30%. The Noto area was over 20% and Kaga area was less than 20%. The lodger number of the spa area near the station remarkably increased due to the beginning of the operation. It is necessary to review the comfortable Omotenshi (hospitality) to the guests and the staffing management in the high increasing areas. It is also necessary to level the fluctuation on the characteristic of the lodger number by holding events. There are the dispersion in the lodger number every month. The leveling leads to the effective utilization of the resources, especially capacity and the meal supply.

In future, it needs to be examined on foreign visitors (inbounds) who have an enhanced economic effect. It will also contribute Japanese globalization. Increasing the inbounds is one of Japanese strategies.

Acknowledgements. A part of the data is provided by Mr. Shinya Inoue (Director of The Yamanaka Onsen Inns Association). The authors gratefully acknowledge to him for his kindness and effective discussions.

References

1. Bureau JT (2013) Viewpoint and practice of the management in tourism spot, Maruzen
2. Insatsu N (2015) White paper on aged society. Cabinet Office (in Japanese)
3. Oyabu T, Nakajima M, Ohe Y, Hosono M (2013) Tourism and regional development. Kaibun-do (in Japanese)
4. Kagaku Sha K (2015) Introduction to tourism informatics. Society for tourism informatics in Japan (in Japanese)
5. Shinbun H (2016) The demand of Obama-Kyoto. Morning edition (in Japanese)
6. Shinbun HC (2015) Facilities user of Ishikawa prefecture becomes the most. Morning edition (in Japanese)
7. Shinbun HC (2015) Tourists exceeded one million. Morning edition (in Japanese)
8. Shinbun HC (2015) Tourists increased remarkably in each spot. Morning edition (in Japanese)
9. Shinbun HC (2016) 0.29 million foreign visitors in Kenroku-en garden. Morning edition (in Japanese)
10. Shinbun HC (2016) The advancement of the achievement (the aim 60,000 visitors). Morning edition (in Japanese)
11. Suda H (2009) Tourism. Gakugei-shuppan (in Japanese)

Charging Infrastructure Allocation for Wireless Charging Transportation System

Min Seok Lee and Young Jae Jang(✉)

Korea Advanced Institute of Science and Technology,
Daejeon, Republic of Korea
yjang@kaist.edu

Abstract. The dynamic wireless charging electric vehicle (DWC-EV) is the most developed form of EVs which use wireless power transfer (WPT) technology. The vehicle can be charged using a wireless charger embedded under the road, and an extremely competitive advantage of DWC-EV is that it can be charged while on the move. Thus, this system allows EVs to overcome the range limitation, long charging time, and heavy battery weights. In DWC-EV system, the battery and charging infrastructures called power tracks are major elements to operate vehicles, and these factors occupy large portion of initial investment cost for establishing the system. This paper suggests a mathematical model to optimally find the battery capacity and the allocation of power tracks in a multiple-route public transportation system. In addition, genetic algorithm (GA), which is one of widely used metaheuristics, is proposed as a solution approach to solve large and complex problems. Since chromosome design determines the format and information of chromosomes, it is crucial to the quality of the solutions and the performance of GA. We suggest two chromosome design methods and the methods to deal with infeasible chromosomes. A number of numerical experiments for comparing solutions and computation time are conducted, and we use the optimal solutions suggested by CPLEX solver as criteria. We also show that the proposed GA design method can be useful and effective compared to the exact solution approach based on the mixed integer programming (MIP).

Keywords: Dynamic wireless charging · Electric vehicle · System optimization · Genetic algorithm

1 Introduction

Electric vehicles (EVs) have attracted international attention as an alternative to internal combustion engine (ICE) vehicles, with the aim of reducing the dependence on petroleum and adverse environmental effects. Many countries and authorities have devoted much research and development to EVs, and numerous policies have also been suggested with respect to electric transportation systems

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_137

[15]. The number of EV users has grown in accordance with the global change, and in particular, the market share of plug-in EVs (PEVs) becomes larger. In addition, wireless power transfer (WPT) technology for EVs has recently been introduced to overcome the disadvantages of the existing PEVs, for example short travel distance and range anxiety.

The dynamic wireless charging electric vehicle (DWC-EV) is the most advanced form of EV, and uses a wireless charger embedded under the road, which charges the vehicle when it moves or stays on it. As it is possible to charge during operation, the driver does not need to stop for a long time to recharge the battery. The battery size can be reduced as the charging infrastructure can supplement the electric power throughout the operation. The online electric vehicle (OLEV) is a DWC-EV developed by the Korea Advanced Institute of Science and Technology (KAIST). In DWC-EV system, however, the charging infrastructure and the battery account for the largest proportion of the initial investment cost. InGumi city, about 80% of the total cost was allocated to these two key factors for setting the system [3,12]. In other words, the investment cost is mainly determined by battery capacity and the allocation of the charging infrastructure, so we focus on these factors in this study. To install a DWC-EV system at a reasonable cost, an optimum level of charging infrastructure and an estimation of battery capacity are required.

The purpose of this study is to identify the optimal allocation of charging infrastructure and battery capacities for the stable and economic operation of the DWC-EV system which is applied to public transportation buses. We focus on the minimization of the initial investment cost, not the operational cost. In this study, we suggest a mixed integer programming (MIP) model considering a multiple-route public transportation system. To find optimal solutions for the MIP model, specific algorithms would typically be used, but obtaining optimal solutions to the complex problems with these algorithms is difficult. In this case, genetic algorithm (GA), a widely used metaheuristic, can be useful for finding optimal or near-optimal solutions within a short time. Therefore, we design a GA procedure to solve the multiple-route problem. We analyze the overall performance of the solutions suggested by GA and also compare it with the optimal solutions obtained by solving MIP problem.

Several studies have been conducted on the system design for EVs, and some focus specifically on infrastructure planning and economic analysis of hybrid EVs or PEVs [4–6,9,14]. These studies generally focus on the deployment of charging stations or determination for the capacity of the charging stations. Also, recent studies related to the system design of DWC-EVs have dealt with the minimization of total cost by determining the battery size and the positions of power transmitters [7,8,8,10,11,13]. These studies have considered single route cases or do not suggest practical methods to solve large and complex problem. In actual, the optimization approach for these cases is not appropriate if the DWC-EV application is planned at the city or multiple district scale.

The remainder of this paper is organized as follows. Section 2 presents the basic characteristics of a DWC-EV system, optimization issues, and several main assumptions used in the optimization model. Section 3 describes the optimization modeling of the DWC-EV system, and Sect. 4 introduces the GA process and chromosome design methods used to suggest solutions for our model. Section 5 solves example problems using GA, and the quality of the solutions are verified by comparing them with the optimal solutions suggested by the CPLEX solver. Section 6 concludes and suggests directions for future study.

2 Description of the DWC-EV System

2.1 The Characteristics of the DWC-EV System

The DWC-EV system has two main parts: the vehicles and the charging infrastructure. The vehicle part mainly consists of a pick-up device for charging, a battery, and a motor for moving vehicles. The charging infrastructure is composed of an inverter and an inductive power cable, which is installed beneath the road. In this paper, we refer to the charging infrastructure as the power track. When a DWC-EV operates on the inductive cable, the power is delivered to the motor or the battery through the regulator. Figure 1 illustrates the main components of the DWC-EV system. The wireless power transfer system is described in detail by Ko and Jang [10] and Ahn et al. [2].

The battery and power tracks are essential elements of the DWC-EV system and constitute most of the total initial investment. It is important to find an appropriate balance between the battery capacity and the allocation of power tracks, but it is not economically viable to install power tracks under whole routes or to equip vehicles with excessively large battery capacity. Two extreme examples can help to explain the relationship between battery and power track. If a vehicle has a large battery that is adequate for operation over the whole route, it does not need to collect electric power during operation, and if it is charged at the base station just before operation there is no problem. The battery will here constitute most of the investment cost. If power tracks are installed under the whole route, the vehicle can travel with a very small battery, and the cost of the power track would be much higher than the battery cost. From the examples, we know that the battery cost and the power track cost have a trade-off relation. Therefore, determining the battery capacity and the allocation of the power tracks should be verified precisely in the commercialization of the DWC-EV system. The purpose of the optimization is to minimize the initial investment cost of the DWC-EV system.

2.2 Assumptions of the DWC-EV System

In this study, we focus on the DWC-EV system applied to public transportation buses. Buses must stop at several stations on each route during operation, and begin and terminate their service at the base station. Before modeling the optimization problems, we set several assumptions, which reduce the complexity of the problems and reflect the practical issues.

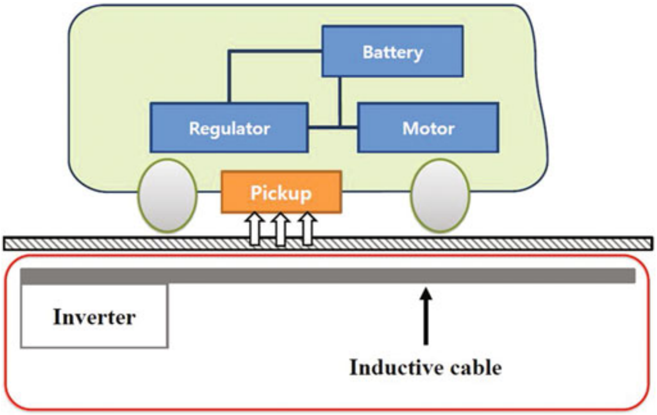


Fig. 1. Major components of DWC-EV

First, the energy consumption between two stations is given. We approximate the consumption is linearly proportional to the distance. The second assumption is that power tracks can be installed only at station areas. In the DWC-EV system, the level of battery charging depends on how long the bus stays on the road where the power track is installed, so clearly installing power tracks at locations where buses stay for a long time is advantageous. Another assumption is that the amount of charging at stations is linearly proportional to the length of power tracks.

The DWC-EV system also includes operational assumptions. First, if several buses travel on an identical route their battery capacities will the same, but battery capacity can differ according to the route. Buses are also assumed to be fully charged at the base station during resting time. We do not consider the installation cost of charging facilities at the base station in the optimization modeling. We will consider only the cost of facilities used during operation.

3 Optimization Modeling

3.1 Decision Variables

We consider multiple routes for public transportation systems in our study. The investment cost will change according to the number and the length of power tracks, so we must identify the number, length, and location of power tracks. Suppose that there are a total of N stations and M routes in the system. We define the integer decision variable X_s as the length of the power track installed at station s $(s = 1, \cdots, N)$. Also, let y_s be the binary variable for whether the power track is installed at station s or not. The battery capacity of a vehicle in route r $(r = 1, \cdots, M)$ is denoted by E_r^c. Figure 2 describes decision variables for a multiple-route optimization model.

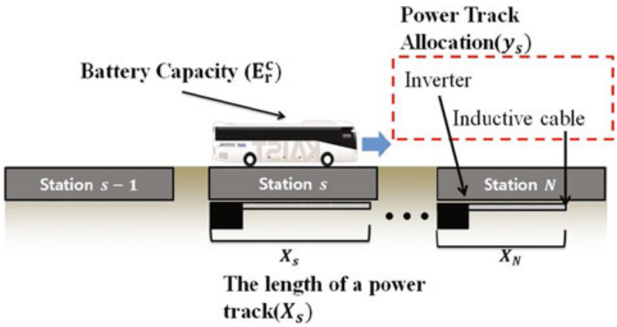


Fig. 2. Decision variables for a multiple-route optimization model

3.2 Objective Function

The objective function is composed of the battery cost and the infrastructure cost. Let c_b be the battery cost per unit capacity. Thus, we can represent the battery cost of a vehicle on route r as $c_b E_r^c$. The power track related cost is composed of a fixed cost and a cable cost. The fixed cost includes inverter installation and setup costs. The length of power tracks has a direct influence on cable cost. The longer the power track is, the more cost is incurred. We define c_f and c_v as the fixed cost and the cable cost, respectively. The power track installation cost at station s is represented by $\sum_{s=1}^{N} c_f \times y_s + \sum_{s=1}^{N} c_v \times X_s$. If there are k_r vehicles on route r, the objective function of the multi-route system is expressed as follows:

$$\sum_{r \in R} c_b \times k_r \times E_r^c + \sum_{s=1}^{N} c_f \times y_s + \sum_{s=1}^{N} c_v \times X_s.$$

3.3 Constraints

In most cases, battery manufacturers set maximum and minimum energy levels for battery stability, so the energy level is recommended to remain within a fixed range. The maximum and minimum levels are represented by multiplying a constant and the battery capacity. These equations are as follows:

$$E_r^u = n^u \times E_r^c,$$
$$E_r^l = n^l \times E_r^c,$$
$$0 \leq n^l \leq n^u \leq 1.$$

The number of stations on route r is denoted by $N(r)$, and k^{th} station on route r is represented by $I(r,k)$. Let $E_{r,k}$ be the energy level of the battery at the moment when the bus leaves k^{th} station on the route r. The battery will not have a higher energy level than E_r^u even if charged for a very long time. Surplus energy will then result, which cannot be stored in the battery. We call this remaining energy excess energy. The excess energy quantity at k^{th} station of route r is denoted by variable $q_{r,k}$. The energy level $E_{r,k}$ is written in the following form:

$$E_{r,k} = E_{r,k-1} - d_{r,k-1} + p \times X_{I(r,k)} - q_{r,k} \quad r \in R, \quad k = 1, \cdots N(r),$$

where $d_{r,k}$ means the energy consumption between k^{th} station and $k + 1^{th}$ station on route r, and p is the amount of energy charge per unit length of the power track.

The system must maintain the battery level within the upper and lower bounds while in operation.

$$\begin{aligned} &E_{r,k} \leq E_r^u \quad r \in R, \quad k = 1, \cdots N(r), \\ &E_{r,k} - d_{I(r,k)} \geq E_r^l \quad r \in R, \quad k = 1, \cdots N(r). \end{aligned}$$

The DWC-EV bus is fully charged when it leaves a base station.

$$E_{r,1} = E_r^c \quad r \in R.$$

The length of the power track cannot exceed the length of the station area. Installing very short power tracks is not economically viable. The following inequality constraint states the upper and lower bounds of the length of the power track.

$$y_s \times X^{\min} \leq X_s \leq y_s \times X^{\max} s \in S,$$

where $X^{\min}$ and $X^{\max}$ are the minimum and the maximum length of the power track.

We develop the following mathematical model for the optimization problem using the objective function and constraints explained above.

$$\min \sum_{r \in R} k_r \times c_b \times E_r^c + \sum_{s \in S} c_v \times X_s + \sum_{s \in S} c_f \times y_s \tag{1}$$

$$\text{s. t. } E_{r,k} = E_{r,k-1} - d_{r,k-1} + p \times X_{I(r,k)} - q_{r,k} \quad r \in R, \quad k = 2, \cdots N(r) \tag{2}$$

$$E_{r,k} \leq E_r^u \quad r \in R, \quad k = 1, \cdots N(r) \tag{3}$$

$$E_{r,k} - d_{I(r,k)} \geq E_r^l \quad r \in R, \quad k = 1, \cdots N(r) \tag{4}$$

$$E_{r,1} = E_r^u \quad r \in R \tag{5}$$

$$y_s \times X^{\min} \leq X_s \leq y_s \times X^{\max} s \in S \tag{6}$$

$$y_s \in \{0, 1\} \quad s \in S \tag{7}$$

$$X_s \text{ and } E_r^c \text{ are zeros or positive intergers } s \in S. \tag{8}$$

4 Genetic Algorithm

Several algorithms are used to solve MIP problems, such as the branch and bound algorithms. However, these algorithms cannot suggest optimal solutions for large-scale problems in a timely manner. In this case, heuristic methods or metaheuristics are more useful for solving the problems. The genetic algorithm (GA) is a widely used metaheuristic, and we use it as our main solution approach in this study. Chromosome design determines the format of chromosomes, and the chromosome's information is crucial to the quality of the solutions and the performance of GA. We suggest two chromosome design methods for our problem.

4.1 Chromosome Design with Battery Capacity

(1) Format of chromosome

The first chromosome is divided into two parts. Figure 3 describes the format of the chromosome, and the first part represents the battery capacities. The value of each gene represents the battery capacity corresponding to each route. The second part of the chromosome expresses the length of the power track at each station. Therefore, this chromosome format contains information about the decision variables E_r^c and X_s.

However, infeasible chromosomes can be produced with this format. Specifically, several chromosomes created during generation will not satisfy the energy dynamics constraint introduced in Eqs. (3) and (4). The infeasibility of the chromosome can be dealt with in two ways; either by giving a large penalty to the chromosome, such as very bad fitness value, or by applying repair mechanism to make it feasible [1]. We deal with these two methods in numerical experiments.

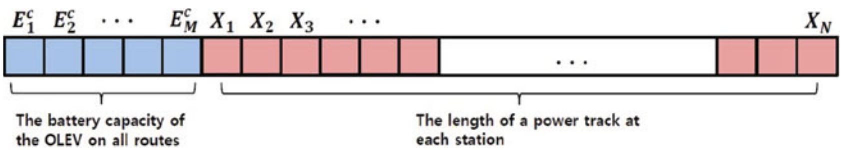


Fig. 3. The structure of a chromosome with genes of battery capacity

(2) Repair mechanism

From a generated chromosome, we calculate the energy level of the battery using Eqs. (2) and (5) at all stations on multiple routes. We suppose that we start the investigation from kth station of route r. We calculate $E_{r,k} - d_{r,k}$, which is defined as remaining energy. From now on, we assume that the lower bound of energy level E_r^l is zero for convenience. If the remaining energy is greater than or equal to zero, we go to the $k+1^{th}$ station and calculate the remaining energy at the $k+2^{th}$ station again. If k^{th} station is the last station of route r, the mechanism starts from the first station of route $r+1$. The mechanism will terminate after the investigation of the last station on the last route.

However, the remaining energy $E_{r,k} - d_{r,k}$ may result, which is less than zero. In this case, we calculate the additional battery capacity and the length of power tracks, respectively, to supplement the shortage of energy. The additional battery capacity or the length of the power track can replenish the energy amount $d_{r,k} - E_{r,k}$, and then we calculate the corresponding cost for each additional element. If the additional battery cost is cheaper than the additional power track cost, we add the battery capacity to the existing battery capacity of route r instead of the power track. In the opposite case, we attach the additional power track to the existing power track. However, if the length after adding the additional power track is longer than the upper bound, the rest of the additional power track is converted to battery capacity. This generated battery capacity is then added to the existing capacity. When the length after adding the additional power track is shorter than the lower bound, we choose additional battery capacity instead of power track.

Next, we modify the energy level $E_{r,k}$, and go to the $k+1^{th}$ station and calculate the remaining energy at the station again. The mechanism should start from the first station of the first route and finish after the process of the last station on the last route. The new chromosome created by the repair mechanism will satisfy all constraints in the MIP model.

4.2 Chromosome Design Without Battery Capacity

(1) Format of chromosome

In this format, we allocate only the power track part of the previous format to a chromosome. Each gene of a chromosome expresses the length of the power track at each station. Therefore, the decision variable X_i in the MIP model corresponds to the i^{th} gene of the chromosome. Figure 4 describes the structure of the chromosome. Although the chromosome does not have information about battery capacity, it is possible to calculate the minimum battery capacity to operate the DWC-EV without depletion of energy based on the gene values of power tracks. Here, there is no infeasible chromosome in terms of energy dynamics constraints, so the repair mechanism is not needed in this format, and we can reduce the waste of battery capacity as we calculate the minimum battery capacity for each chromosome. However, we must calculate the battery capacities for all chromosomes created until the last generation, so more calculation time will be spent on the GA procedure than the previous format.

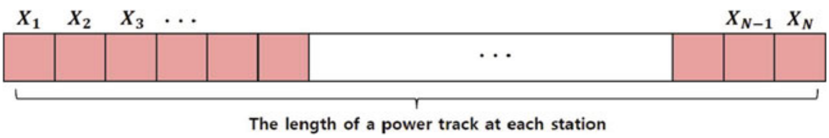


Fig. 4. The structure of a chromosome without genes of battery capacity

(2) Calculation of battery capacity

Let $W_{r,k}$ be the required energy level of the battery at the moment when the bus arrives at k^{th} station on route r. We assume that $W_{r,N(r)+1}$ is zero, which is the lower bound of the energy level, because station N(r)+1 represents the base station of route r. When the bus departs from $N(r)^{th}$ station on route r, it needs at least $d_{r,N(r)} + W_{r,N(r)+1}$. If a power track of $X_{I(r,N(r))}$ in length is installed at $N(r)^{th}$ station, $W_{r,N(r)}$ would be max $0, d_{r,k} + W_{r,k+1} - p \times X_{I(r,k)}$. In this manner, it is possible to explain other $W_{r,k}$. Therefore, $W_{r,k}$ can be expressed as follows:

$$W_{r,k} = \max\left(0, d_{r,k} + W_{r,k+1} - p \times X_{I(r,k)}\right), r \in R, k = 1, \cdots N(r) \tag{9}$$

To calculate the battery capacity E_r^c on route r, we define $bat(k)$ as the required battery capacity to move from k^{th} station to $k+1^{th}$ station on route r. The battery will consume $d_{r,k}$ until arriving the next station, and the energy state at $k+1^{th}$ station should be $W_{r,k+1}$. Thus, $bat_r(k)$ is the sum of $d_{r,k}$ and $W_{r,k+1}$.

$$bat_r(k) = d_{r,k} + W_{r,k=1}, r \in R, k = 1, \cdots N(r). \tag{10}$$

The battery capacity E_r^c is the maximum value among all $bat_r(k)$ to cover all movements between two stations. Therefore, E_r^c is represented as follows:

$$E_r^c = \max_{k=1,\cdots N(r)} (bat_r(k)), r \in R. \tag{11}$$

4.3 GA Settings and Procedures

Fitness function is an indicator of the quality of the solutions in GA. The objective function of the mathematical model is the fitness function of our problem. The purpose of the GA procedure is to find a solution that minimizes the fitness value. Before starting GA procedure, we generate the initial population to enhance the GA performance. The chromosome in the initial population is generated using the optimization results of each single-route problem. Most single route solutions are easily obtained using the algorithm for MIP model.

Selection function finds parents for the reproduction of new chromosomes using the fitness value. In our case, the selection function is based on the roulette wheel selection. After that, a portion of offspring is reproduced by the crossover function, which generates a new chromosome by combining a part of each parent. First, the crossover function creates a random binary vector that has a length identical to a chromosome. From the first parent, genes placed at the location 1 in the binary vector are selected. Those placed at the location 0 are selected from the second parent. The selected genes from the two parents generate a new chromosome.

The mutation function prevents the solution from converging to a local optimum, and is used to find various solutions in the searching space. In our case, it generates a random vector with length the same as that of the chromosome, and the random vector is added to a selected chromosome. If the new chromosome satisfies the bounds and constraints, it enters the next generation, and we can call it a mutated chromosome. We also set the elite reproduction which means that some of outstanding chromosomes in terms of fitness value must survive to the next generation.

In general, GA stops the procedure when several conditions satisfy the stopping criteria. In our case, the number of generations is selected as a stopping criterion. When the algorithm carries out the last generation, GA will find and suggest the best solution in that final generation.

5 Numerical Experiments for a Multiple-Route DWC-EV System

In this section, numerical experiments for a multiple-route model are handled, and solution approaches introduced in the previous sections are used to solve the problem. We consider 120 stations in Gumi city, which are used for three real routes, 1-1, 8, and 17, in the public transportation system. We add three hypothetical routes to these existing routes, which also use some parts of the 120 stations. Thus, there are 6 routes in the example problem. We assume that the energy consumption between two stations is constant, and the data is already given. Several numerical experiments are conducted as the number of the routes increases from 1 to 6. From routes 1 to 6, there are 47, 32, 46, 36, 36, and 36 stations, respectively, and some of those routes share the same stations.

Table 1. The comparison of average computational time (in seconds)

Description	Values
Unit battery cost (c_b)	800
Fixed cost (c_f)	20,000
Unit cable cost (c_v)	100
Charging rate (p)	0.5
The number of buses on each route (k_r)	5
The maximum length of the power track ($X^{\max}$)	30
The minimum length of the power track ($X^{\min}$)	10

The MIP model for the DWC-EV system is solved using CPLEX 12.5 embedded in a JAVA environment, and the GA coded with the MATLAB Global Optimization Toolbox is used to solve the problems. There are three approaches in the GA. First, we apply the chromosome format with battery capacity, which does not conduct a repair mechanism. This approach is defined as GA method 1. In this method, infeasible chromosomes should have infinite fitness values. The second approach is GA method 2, and carries out the repair mechanism to the same format. The last approach, GA method 3, uses the chromosome format without the battery capacity. Other GA procedures are applied identically to all approaches. The constant values for the numerical experiments are summarized

Table 2. Parameter settings for GA

Description	Values
Population size	1,000
The number of generations	1,000
Crossover fraction	0.84
Mutation fraction	0.01
Elite fraction	0.15

in Table 1. In this example, the maximum and the minimum energy level of the battery are the battery capacity E_r^c and 0, respectively. Several parameter settings for the GA procedures are defined in Table 2. These values are determined from a number of test experiments to find the better solutions.

We select the routes in a regular sequence starting from route 1; for example, if we select 3 routes, we choose route 1, 2, and 3. The total cost obtained from CPLEX and three GA approaches is listed in Table 3 according to the number of routes.

Table 3. The comparison of total cost between solution approaches

Algorithm	CPLEX	GA method 1		GA method 2		GA method 3	
Description	Obj. value	Best	Average	Best	Average	Best	Average
One route	325,100	325,100	357,139	325,100	345,759	325,100	326,934
Two routes	611,900	623,300	680,141	614,900	640,350	611,900	612,215
Three routes	967,400	1,031,200	1,098,603	982,900	1,032,737	975,100	977,649
Four routes	1,074,000	1,161,500	1,225,973	1,117,300	1,154,917	1,105,200	1,108,030
Five routes	1,266,600	1,476,900	1,554,957	1,373,500	1,446,741	1,314,500	1,321,851
Six routes	1,366,000	1,595,400	1,684,494	1,471,400	1,535,578	1,416,400	1,438,070

The GA results of Table 3 are obtained from 100 experiments for each case. The results of CPLEX are objective values, which are calculated by optimal solutions, so it is impossible to have costs lower than the results. These values can be used as criteria to assess the performance of the GA results. In the case of one route, the best value for each GA method is identical to the objective value, but average values are different from each other. GA method 3 suggests the smallest average value of all GA methods, while GA method 1 has the worst average value. In the case of two routes, the best value of GA method 3 also has the same value as the objective value, but GA methods 1 and 2 do not suggest the objective value, and GA method 3 still draws the best values between the three methods in both the best and average values. From the case of three to six routes, no GA methods suggest optimal results, and the performance of solutions in GA method 3 is always better than the others, in the same manner.

GA method 2 takes second place in every case, and GA method 1 proposes the worst results.

Though methods 1 and 2 use the same format of chromosome in the GA procedure, there are huge differences in the results, depending on whether the repair mechanism is used or not. In GA method 1, many infeasible chromosomes can be created during the 1,000 generations, so feasible chromosomes that do not have good fitness value in comparison with the optimal results will easily survive. In addition, the speed of improvement in the chromosomes slows, as the population generally consists of poor chromosomes. The quality of the best chromosome in the final generation will consequently be worse than in other methods. However, if the repair mechanism is applied to GA, it is impossible to have infeasible chromosomes, so all chromosomes can have specific fitness values that are not infinite. Unlike GA method 1, in which feasible chromosomes compete with many infeasible chromosomes, GA method 2 creates competition between comparable chromosomes, and the outstanding chromosomes will survive. Some infeasible chromosomes may also be changed to remarkable chromosomes during the repair mechanism. Therefore, we claim that GA method 2 can suggest better solutions than GA method 1.

GA method 3 provides the best solutions among all GA methods because of the format of chromosomes, and the fact that battery capacity plays an important role in the total cost. Figure 5 represents the total battery cost of the best chromosomes in all GA methods. GA method 3 suggests the smallest battery cost in all cases except for the one route case. GA method 3 can have an advantage in the battery cost as the battery capacities in the chromosome format are calculated to have the minimum values without depletion when it is driving. In contrast, other methods randomly allocate battery capacities, regardless of the gene values in the power track part.

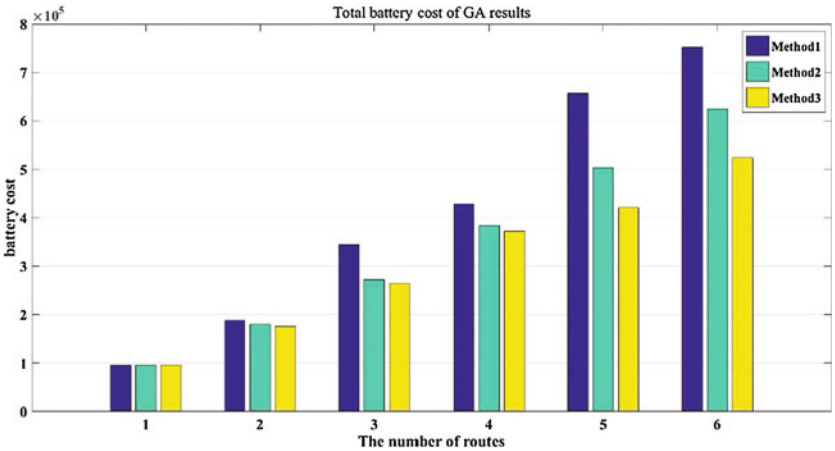


Fig. 5. A comparison of total battery cost in the GA methods

The power track costs of the best chromosomes in all GA methods are shown in Fig. 6. In contrast with the battery cost, the power track cost has the biggest value in GA method 3. The power track costs are identical in the different methods for the one route case, in the fourth case GA method 3 has a slightly smaller value, but for the other cases method 3 has the largest. From the perspective of a trade-off relationship, the results are reasonable as a small amount of power track is allocated to the solution with large battery capacity, and a large amount of power track is provided to the solution with small battery capacity. By comparing battery and power track costs, we confirm that the difference in the battery cost has a huge effect on the trend of the total cost.

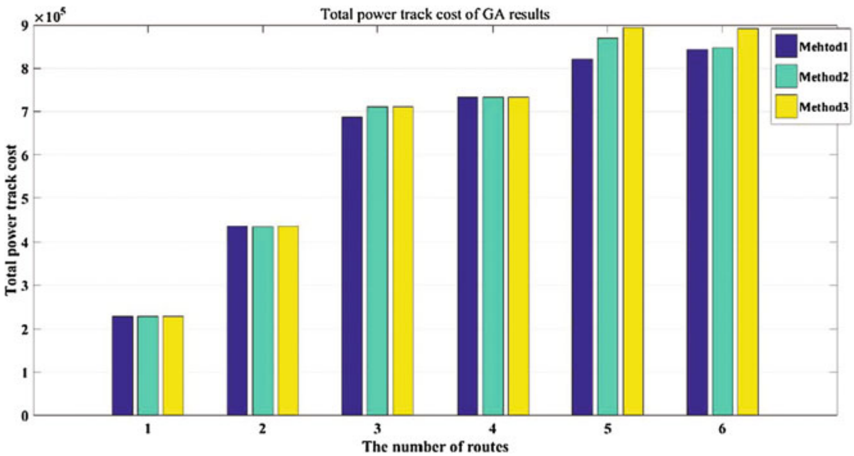


Fig. 6. The comparison of total track cost in GA methods

A comparison of the average calculation time in CPLEX and all GA methods is given in Table 4. The calculation time of GA method 1 and GA method 2 very slightly increases as the number of routes increases. The results of GA method 2 are slightly higher than those of GA method 1 because of the repair mechanism. The repair mechanism does not take up much time during the GA procedure as not all chromosomes are infeasible, and the modification is carried out for only for several genes in the infeasible chromosome. However, the results of GA method 3 indicate that the calculation time for each case is longer than for methods 1 and 2. The slope of the graph is also the steepest of the three methods. In the case of six routes, GA method 3 takes 381.15 seconds to suggest a solution, a calculation time about 5 times that of methods 1 and 2, as it must calculate the minimum battery capacity for all chromosomes in every generation, and the tasks for calculating the battery capacity also increase with the increase in the number of routes.

In the first and second cases, CPLEX is faster than all GA methods, but the time then increases exponentially from the third case, and for the six routes it

Table 4. The comparison of average computational time (in seconds)

Algorithm	CPLEX	GA method 1	GA method 2	GA method 3
One route	0.74	70.62	70.96	110.78
Two routes	1.26	71.4	72.14	156.05
Three routes	409.89	72.21	75.21	227.67
Four routes	619.68	77.13	80.36	274.96
Five routes	902.86	78.98	81.13	321.9
Six routes	40828.58	80.96	83.86	381.15

takes over 11 hours. This trend shows the characteristics of the exact algorithm, such as the Branch and Bound algorithm, which means that the exact algorithm is more sensitive to the size of problems compared to the GA. From the numerical experiments, we identify that GA methods 1 and 2 have an advantage in calculation time compared to GA method 3, but the quality of solutions from method 3 outperforms the other two. The GA can suggest reasonable solutions in a timely manner, but does not always guarantee optimal solutions. Therefore, the GA can be useful for large scale problems.

6 Conclusion

This study has introduced a new kind of dynamic wireless charging electric vehicle (DWC-EV) system. A mixed integer programming (MIP) model, which reflects multiple-route public transportation systems, was suggested to identify the economic allocations of the charging infrastructure and the battery capacity of the vehicle. A GA was introduced as a solution approach, and two approaches for chromosome design were suggested. Numerical experiments for an exact algorithm and a GA were conducted, and we analyzed the performance of three GA methods by comparing the optimal results of CPLEX. We checked that the repair mechanism for infeasible chromosomes improved the algorithm performance. A chromosome design without battery capacity was found to provide better solutions than other GA methods, though its calculation time was slow. These solutions are thus almost optimal solutions.

In this study, several assumptions were applied to develop the mathematical model. The assumption that the energy consumption between two stations is linearly proportional to the distance is particularly impractical, as this does not reflect the uncertainties of traffic or unpredictable situations. The assumption that the amount of charging is linearly proportional to the length of the power track was used to simplify the model. These assumptions should be proved quantitatively, and robust optimization can be used to consider uncertainties. The procedure of the GA must be established to solve the problem efficiently and effectively. After extending the limits of this study, we plan to apply our approach to broader multiple-route systems on a city or national scale. The GA is a simple approach and can be useful for discovering solution properties, and the solutions obtained can be reference points for heuristics and other algorithms, which will be examined in future research.

References

1. Abdelmaguid TF, Dessouky MM (2006) A genetic algorithm approach to the integrated inventory-distribution problem. Int J Prod Res 44(21):4445–4464
2. Ahn S, Suh NP, Cho DH (2013) Charging up the road. IEEE Spectr 50(4):48–54
3. Bulletin M (2013) Korean electric buses go wireless. Wall Street J 32:161–174
4. Chung SH, Kwon C (2015) Multi-period planning for electric car charging station locations: a case of korean expressways. Eur J Oper Res 242(2):677–687
5. Dong J, Liu C, Lin Z (2014) Charging infrastructure planning for promoting battery electric vehicles: an activity-based approach using multiday travel data. Transp Res Part C Emerg Technol 38(1):44–55
6. He F, Wu D et al (2013) Optimal deployment of public charging stations for plug-in hybrid electric vehicles. Transp Res Part B Meth 47(1):87C101
7. Jang YJ, Jeong S, Ko YD (2015) System optimization of the on-line electric vehicle operating in a closed environment. Comput Ind Eng 80(C):222–235
8. Jang YJ, Suh ES, Kim JW (2015) System architecture and mathematical models of electric transit bus system utilizing wireless power transfer technology. IEEE Syst J 99(2):1–12
9. Jia L, Hu Z et al (2012) Optimal siting and sizing of electric vehicle charging stations. Dianli Xitong Zidonghua/Autom Electr Power Syst 36(3):54–59
10. Ko YD, Jang YJ (2013) The optimal system design of the online electric vehicle utilizing wireless power transmission technology. IEEE Trans Intell Transp Syst 14(3):1255–1265
11. Ko YD, Jang YJ, Min SL (2015) The optimal economic design of the wireless powered intelligent transportation system using genetic algorithm considering nonlinear cost function. Comput Ind Eng 89(C):67–79
12. Shin J, Shin S et al (2014) Design and implementation of shaped magnetic-resonance-based wireless power transfer system for roadway-powered moving electric vehicles. IEEE Trans Ind Electron 61(3):1179–1192
13. Song YU, Park S et al (2013) A study on the validity of the infrastructure construction cost for the commercialization of online electric vehicles. J Soc e-Bus Stud 18(1):218–222
14. Sweda T, Klabjan D (2011) An agent-based decision support system for electric vehicle charging infrastructure deployment. In: Vehicle Power and Propulsion Conference, pp 1–5
15. Tali T, Telleen P (2013) Global EV outlook: understanding the electric vehicle landscape to 2020. Int Energy Agency 87:22–30

The Research of Incentive Model Based on Principal-Agent for R&D Personnel in System-Transformed Institutes

Ying Wei(✉) and Jun Gang

Sichuan Institute of Building Research, Chengdu, People's Republic of China
452758107@qq.com

Abstract. The lack of innovation and vitality from R&D personnel is a major problem to system -transformed institutes' innovation activities, and the short of rational and effective incentive mechanism is the source. Based on the analysis of present compensation and motivation, this paper established an incentive model of income distribution and a bonus incentive model by principal-agent theory, and then discussed the advantages and disadvantages of the two incentive models. The result shows that the incentive model of income distribution is more conducive to stimulate the enthusiasm of R&D personnel for system-transformed institutes.

Keywords: System-transformed institutes · R&D personnel · Incentive mechanism · Principal-agent

1 Introduction

Over 10 years' reform, some scientific research institutes have been transformed into enterprises. In this process, the system-transformed institutes actively cultivated the ability of technological innovation with excellent talent teams, and obtained advanced scientific and technological innovation constantly, those are inseparable with the creation and contribution of scientific and technological innovation talents. R&D personnel as the key power of the system-transformed institutes, is irreplaceable in the development and growth of enterprises. Therefore, how to provide maximize innovative space and their creative enthusiasm has become the most important issue for the institutes. But due to historical and practical reasons, there are still some problems such as the inflexible incentive mechanism, the single incentive means and so on, which restrict the innovation ability and enthusiasm seriously. Consequently, whether from the development of institutes itself or from the realization of the personnel value, it is very necessary to study the motivation of R&D personnel in system-transformed institutes.

Although some institutes have already developed incentive policies, but most of them are one-time awards, and usually pay attention to the number of achievements while ignore the quality and the efficiency, which are not conducive to

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_138

long-term enthusiasm for the R&D personnel. Many scholars at home and abroad have carried out a lot of research work about the motivation. Ou [5] researched on incentive of science and technology plan project by principal agent, analyzing the incentive factors and expound the design ideas of project incentive mechanism. Dikolli and Kulp [1] studied the impact of interaction between performance indicators of each task with multitask principal agent model. From the perspective of creating core competence, Gao [3] focused on human resources management problems and disadvantages about system-transformed institutes. Guo and Zhang [4] summarized the policies and regulations of the income distribution of personnel achievements, and proposed suggestions to improve the incentive mechanism. Ding and Chen [2] analyzed the Chinese S&T personnel motivation legal laws and regulations, and put forward policy recommendations to break the barriers to strengthen the transformation achievements motivation.

In summary, the current researches mainly concentrate on the demand characteristic, incentive factors and other aspects of R&D personnel, having certain referential significance to solve the motivation problem. On the basis of above researches, combining with the actual situation of system-transformed institutes, this paper took use of the principal-agent theory, optimizing the incentive mechanism design, discussed the incentive model of income distribution and bonus incentive model respectively though the balance of benefits and risks among the R&D personnel.

2 Background

2.1 The Connotation of the Collaborative Governance

A questionnaire survey for 18 system-transformed institutes in Sichuan has been carried out. As we can see in Fig. 1 that R&D personnel are always with high degree and different titles, showing that knowledge talents are the backbone of personnel team. Speaking of the salary system and incentive mechanism, we find that there are two main problems in current R&D personnel innovation. One is how to combine the performance appraisal with income, the other is the relationship between achievements and income. Each institute is actively exploring or thinking to solve the innovation problems. About 63.6% of institutes have a preliminary study on how to establish the performance evaluation and incentive mechanism, while others do not give a complete management advice on scientific and technological performance appraisal, still in a state of wait and see on the issue of income distribution about achievements.

The composition of R&D personnel income includes intrinsic compensation and extrinsic compensation, which can be seen in Fig. 2. Further, the intrinsic compensation is closely related to incentives, containing basic salary, merit pay, research prize and others. And the R&D personnel's income is mainly dependent on the basic salary and merit pay, taking 81.4% of the whole income, while less incentive income in Fig. 3. Most system-transformed institutes have not established a performance evaluation system different from the business production personnel, especially the case of performance appraisal system is not well sound.

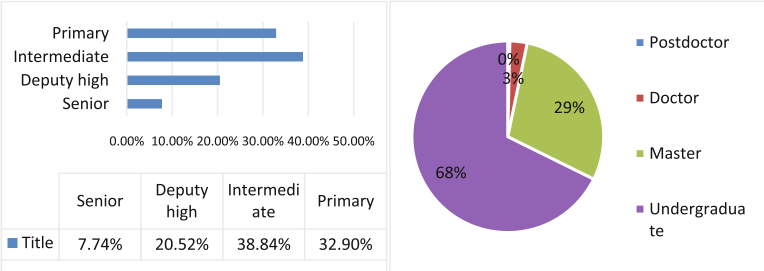


Fig. 1. The composition of R&D personnel in 18 institutes

Even though some institutes set up a merit pay, but not link to the performance of scientific research, so the income systems are still egalitarianism with small incentive income, especially lack of long-term incentive mechanism.

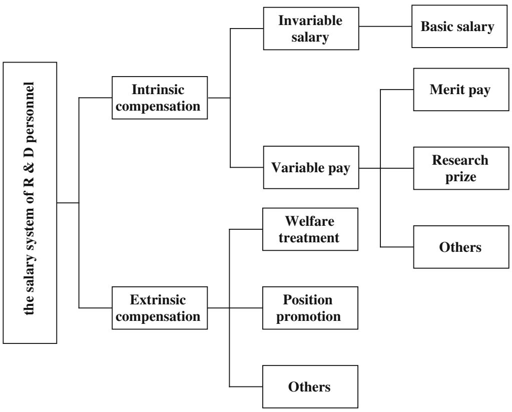


Fig. 2. The salary system of R&D personnel

3 Incentive Modeling

3.1 Income Distribution Incentive Model

Through the investigation and analysis above, it is not difficult to find that even there are four parts of R&D personnel income, the first two are actually relatively fixed, which can be seen as fixed income. And the other two are more fluctuant and associate with the degree of effort, can be regarded as incentive income. Assuming that the effort degree is proportional to the output, so the output of R&D personnel can be shown as:

$$X = W_0 + a + \varepsilon \quad \varepsilon \sim N(0, \sigma). \tag{1}$$

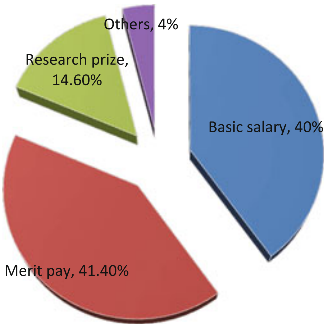


Fig. 3. The composition of intrinsic compensation

At the same degree of effort, the output embodied in two levels, the single value of scientific and technological achievement and the number of achievement. For the labor of R&D personnel is the knowledge labor, long term on a small number of achievements will get more value than those focus on numbers of achievements. For example, a patent will be specialized or in 10 patents be invented in one year. Obviously, the 10 patents tend not to produce too much real value. Therefore, hypothesis that and satisfy the following relations:

$$a = qv^{1/n}. \tag{2}$$

So, the actual benefits from R&D personnel to the institutes can be expressed as:

$$S = vq = v^{\frac{n-1}{n}}a. \tag{3}$$

Assume that R&D personnel share the results of income in accordance with a certain proportion k_1, so the personal income can be expressed as:

$$I = W_0 + v^{\frac{n-1}{n}}a. \tag{4}$$

In addition, R&D personnel have to pay a certain cost as well as the effort. Taking use of the common quadratic model in principal-agent theory, use to represent the effort cost, which is the relationship between effort and cost, so the cost can be expressed as:

$$c_1 = \frac{1}{2}ba^2. \tag{5}$$

Otherwise, comparing with the scientific research institutes, R&D personnel usually take risk aversion which needs to pay a certain risk costs. Set c_2 on behalf of the risk aversion, using Arrow-Pratt measure of absolute risk aversion to construct the risk cost function as follows:

$$c_2 = \frac{1}{2}\rho k_1^2 v^{\frac{2n-2}{n}}\sigma^2. \tag{6}$$

So the actual expected income can be expressed as:

$$\mathrm{E[I]} = W_0 + k_1 v^{\frac{n-1}{n}} a - \frac{1}{2} b a^2 - \frac{1}{2} \rho k_1^2 v^{\frac{2n-2}{n}} \sigma^2. \tag{7}$$

Certainly, most of the institutes usually demand the minimum number of achievements, which record as $\underline{v}$. And set $\underline{q}$ as the lower limit of the single value, so the value of a single scientific and technological achievement shall meet the following constraints:

$$\underline{v} \leq v \leq a/\underline{q}. \tag{8}$$

In summary, the R&D personnel decision-making model is:

$$\begin{aligned} &\max_{a,q} \quad \mathrm{E[I]} = W_0 + k_1 v^{\frac{n-1}{n}} a - \tfrac{1}{2} b a^2 - \tfrac{1}{2} \rho k_1^2 v^{\frac{2n-2}{n}} \sigma^2 \\ &\text{s.t.} \quad \underline{v} \leq v \leq a/\underline{q}. \end{aligned} \tag{9}$$

Similarly, the income of institutes is $(1-k_1)v^{\frac{n-1}{n}} a - W_0$. The risk of institutes is often neutral relative to R&D personnel, pursue the maximization of expected return usually. In addition, R&D personnel usually set a psychological bottom line of personal income, which institutes must take it into consideration. So the decision model of institutes is:

$$\begin{aligned} &\max_{k} \quad \mathrm{E[V]} = (1-k_1) v^{\frac{n-1}{n}} a - W_0 \\ &\text{s.t.} \quad W_0 + k_1 v^{\frac{n-1}{n}} a - \tfrac{1}{2} b a^2 - \tfrac{1}{2} \rho k_1^2 v^{\frac{2n-2}{n}} \sigma^2 \geq I_{\min}. \end{aligned} \tag{10}$$

To sum up, the incentive model under the income distribution model for R&D personnel, the decision model can be shown as:

$$\begin{aligned} &\max_{k} \quad \mathrm{E[V]} = (1-k_1) v^{\frac{n-1}{n}} a - W_0 \\ &\text{s.t.} \quad \begin{cases} W_0 + k_1 v^{\frac{n-1}{n}} a - \tfrac{1}{2} b a^2 - \tfrac{1}{2} \rho k_1^2 v^{\frac{2n-2}{n}} \sigma^2 \geq I_{\min}, \\ \max\limits_{a,q} \mathrm{E[I]} = W_0 + k_1 v^{\frac{n-1}{n}} a - \tfrac{1}{2} b a^2 - \tfrac{1}{2} \rho k_1^2 v^{\frac{2n-2}{n}} \sigma^2, \\ \text{s.t.} \quad \underline{v} \leq v \leq a/\underline{q}. \end{cases} \end{aligned} \tag{11}$$

For the income is directly linked to the transfer profit, so the R&D personnel will try to create the greatest value of achievements under the premise of the minimum completion number. The effort level of R&D personnel is positively correlated with the income coefficient, the greater proportion of income distribution, the more initiative from R&D personnel. Strives for the partial derivatives can get the following results:

$$\frac{\partial k_1}{\partial \rho} < 0, \frac{\partial a}{\partial \rho} < 0, \frac{\partial E[V]}{\partial \rho} < 0, \frac{\partial k_1}{\partial b} < 0, \frac{\partial a}{\partial b} < 0, \frac{\partial E[V]}{\partial b} < 0, \frac{\partial k_1}{\partial \underline{q}} < 0, \frac{\partial a}{\partial \underline{q}} < 0, \frac{\partial E[V]}{\partial \underline{q}} < 0.$$

Thus it can be seen that, in the optimal state, the income coefficient, the effort level and the income of institutes are negatively correlated with the risk aversion, the effort cost and the minimum number of achievements. This suggests that under the incentive model of income distribution, the system-transformed institutes should reduce the limit number of achievements can not only encourage R&D personnel in a greater extent, but also improve the actual income of institutes.

3.2 Bonus Incentive Model

This part will discuss the incentive problem of R&D personnel under the bonus incentive model. Hypothesis that the bonus is positively with the number of achievements, and the achievement bonus of the institute is k_2. Still, the single value v and the number q satisfy the following relations at the same level of effort:

$$a = qv^{1/n}. \tag{12}$$

Then, the income of R&D personnel can be expressed as:

$$I = W_0 + k_2 q = W_0 + k_2 v^{\frac{1}{n}} a. \tag{13}$$

For the bonus is positively with the achievement number, so under the same level of effort, R&D personnel must ignore the quality of the results, and pay more attention to the number of results. Set $\underline{v}$ as the lower limit of the single value of achievements, and the upper limit is $a/\underline{q}$ when meet the minimum number of results, so taking the cost and risk into account, the expected return model of R&D personnel is:

$$\begin{aligned} &\max_{a,v}\ \mathrm{E[I]} = \mathrm{W}_0 + k_2 v^{-\frac{1}{n}} \mathrm{a} - \tfrac{1}{2} b a^2 - \tfrac{1}{2} \rho k_2^2 \mathrm{v}^{-\frac{2}{\mathrm{n}}} \sigma^2 \\ &\text{s.t. } \underline{v} \le v \le a/\underline{q}. \end{aligned} \tag{14}$$

The expected return of institute is:

$$E[V] = S - W_0 - k_2 q = vq - k_2 v^{\frac{1}{n}} a - W_0 = v^{\frac{n+1}{n}} a - k_2 v^{\frac{1}{n}} a - W_0. \tag{15}$$

So, under the bonus incentive model the incentive model can be shown as follows:

$$\begin{aligned} &\max_{k_2}\ \mathrm{E[V]} = v^{\frac{n+1}{n}} a - k_2 v^{\frac{1}{n}} a - W_0 \\ &\text{s.t.}\quad \begin{cases} \mathrm{W}_0 + k_2 v^{-\frac{1}{n}} \mathrm{a} - \tfrac{1}{2} b a^2 - \tfrac{1}{2} \rho k_2^2 \mathrm{v}^{-\frac{2}{\mathrm{n}}} \sigma^2 \ge I_{\min}, \\ \max\limits_{a,v}\ \mathrm{E[I]} = \mathrm{W}_0 + k_2 v^{-\frac{1}{n}} \mathrm{a} - \tfrac{1}{2} b a^2 - \tfrac{1}{2} \rho k_2^2 \mathrm{v}^{-\frac{2}{\mathrm{n}}} \sigma^2, \\ \text{s.t. } \underline{v} \le v \le a/\underline{q}. \end{cases} \end{aligned} \tag{16}$$

There we can get the following conclusions, (1) under the bonus incentive model, in order to maximize the personal interests, R&D personnel will ignore the actual value of the research results and blindly pursue the numbers. As a principal, institutes can set the appropriate technology bonus by assessing the lowest value of achievements. (2) For the bonus amount is proportional to the lowest value, so the lower value of achievements, the more achievements R&D personnel create in the same level of effort, so as to obtain more revenue. (3) Stand in the perspective of the institutes, to avoid the emergence of low value achievements, they should set a threshold of achievements, and strictly evaluate them.

4 Comparison Analysis

Compare the income distribution incentive model with the bonus incentive model (see in Table 1), we can find that: (1) R&D personnel have different pursuit in the two models. R&D personnel pay more attention on the value of achievement under the income distribution incentive model, while pursue to create a great number of achievements in the bonus incentive model. For the system-transformed institutes take profit as the goal, thus using income distribution incentive model more. And the colleges use bonus incentive model frequently when facing with a large number of assessment indicators. (2) In the income distribution incentive model, the revenue sharing factor of R&D personnel is relative with effort cost, risk aversion coefficient and output random error, but has nothing to do with the result itself. While the bonus factor in addition to the impact above, also related to the lowest value of scientific and technological achievements.

Table 1. The comparison of two models

Parameter	Income distribution incentive model	Bonus incentive model
k	$\frac{1}{1+\rho b\sigma^2}$	$\frac{\underline{v}^{\frac{n+2}{n}}}{1+b\rho\sigma^2}$
v	$\frac{1}{b^n \underline{q}^n (1+\rho b\sigma^2)}$	$\underline{v}$
a	$\frac{\underline{q}^{1-n}}{b^n (1+\rho b\sigma^2)^n}$	$\frac{\underline{v}^{\frac{n+1}{n}}}{b+\rho b^2\sigma^2}$
$E[V]$	$\frac{\underline{q}^{2-2n}}{2b^{2n-1}(1+\rho b\sigma^2)^{2n-1}} - I_{\min}$	$\frac{\underline{v}^{\frac{2n+2}{n}}}{2b(1+\rho b\sigma^2)} - I_{\min}$

5 Conclusion

Scientific incentives can attract and retain excellent researchers for scientific research institutes, meeting the ultimate needs of beneficiaries constantly. And salary is the prerequisite to ensure the material needs of R&D personnel, also is the prerequisite and basis for the scientific and technological talents to pursue higher level needs. Even though China's policies and regulations for the allocation of R&D personnel incentive requirements are put forward, most of the institutes only pay fixed income when they can observe the effort level from personnel. Normally, it is difficult to observe the effort level, still let researchers share part of the output share, otherwise R&D personnel will not work hard with basic salaries, which is detrimental to the development of enterprises, also the country. From the analysis, R&D personnel focus on the value is applied to the system-transformed institutes, while a great number of achievements will be made to the colleges. The revenue sharing factor of R&D personnel is relative with effort cost, risk aversion coefficient and output random error, but has nothing to do with the result itself. Therefore, the greater effort cost, the smaller share of output, and the smaller risk for those who fear of hard work. Conversely, the stronger ability, the lower effort cost.

Acknowledgements. This research was conducted with the support of science plan program of Chengdu in 2016 (No. 2015-RK00-00062-ZF).

References

1. Dikolli SS, Hofmann C, Kulp SL et al (2009) Interrelated performance measures, interactive effort, and incentive weights. J Manag Acc Res 21(1):125–149
2. Ding M, Chen B (2016) Research on the strengthening the S&T personnels transformation achievements motivation based on the improvement of the reward mechanism determined by knowledge, technology, management, skills and other factors market. Sci Manag Res 142(3):96–100
3. Gao J (2013) The incentive mechanism of human resources management of scientific research institution after ownership transformation. Municipal Eng Technol 61(2):46–54
4. Guo Y, Zhang S (2015) The incentive mechanism of the income distribution of scientific and technological achievements of technical personnel. Sci Sci Manag S&T 13:183–190
5. Ouyang J, Yan H, Huang C et al (2008) Research on incentive of science and technology plan project based on principal agent. Sci Technol Manag Res 28(8):246–248

Research on the Influencing Factors of Public Sense of Security and Countermeasures in Major Emergencies–An Empirical Analysis on the Public Sense of Security in View Of the Explosion in Binhai New Area in Tianjin

Jing Yang(✉) and Wei Xu

School of Political Science and Public Administration,
University of Electronic Science and Technology of China,
Chengdu, The People's Republic of China
yangjing8056@uestc.edu.cn

Abstract. In the crisis, the biggest psychological threat to the public is the loss of public sense of security. Through the related research, it is hypothesized that critical events, countermeasures of the government and the media, emergency response capability of the individual and the emergency psychological behavior of the group are the four main factors influencing public sense of security in the emergencies. A total of 330 samples were randomly selected from those people who had experienced the explosion in Tianjin. By exploratory factor analysis and the structural equation model examination, the fourfactors all have different degrees of explanation towards the loss of public sense of security. According to the conclusion, we can see that it is the key for the government to construct the governance mechanism by enhancing the emergency response capability of the group improve their certain sense of crisis and their sense of control of the situation.

Keywords: Public sense of security · Critical events · Countermeasures of the government and the media · Emergency response of the individual and the group · Structural equation model

1 Introduction

A crisis is often caused by both the "objective crisis of the events" and "subjective crisis of the individual", but the negative impact of the subjective crisis of individual is often greater than that of the events. Just as the attack on September 11 has happened for 15 years while its impact and psychological fears left to people have a long-term influence on American society. In 2011, Fukushima nuclear accident caused "salt-buying panic" in China, the reason of which is

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_139

people's fear for the crisis. When the crisis comes, the biggest threat to the public is the loss of their sense of security which is a prominent feature of public emergency psychology after the crisis. But why a few events that are actually less harmful are more influential on the public than those that have fatal consequences? Why do crisis of the same types have different effects on the public sense of security in a different time and place? This series of events have shown that the formation of public sense of security requires both objective conditions and also subjective factors.

Based on the existing research on the public sense of security, this paper designs a model of the influencing factors of public sense of security and also analyzes the data through questionnaire and the relationship between data and their characteristics based on SPAA as well as the influencing factors of the public sense of security based on the structural equation of AMOS. Besides, it measure how these factors influence the public sense of security, which contributes to the design of the emergency psychological governance mechanism in the emergency.

2 Conceptual Definition and Literature Review

2.1 Definition of the Public Sense of Security

According to scholars at home and abroad, sense of security is referred to the sense of belonging in a community and the sense of self [3]. The "public sense of security" in this article involves the sense of belonging, the certain sense of control and the satisfaction of security needs that the public experience in a certain public environment. Among them, the sense of belonging refers to the warm and the reliable subjective experience of the cares and comfort from other members in the community. The certain sense of control refers to the certain cognitive and controlled experience for the causes, processes and trends of the public crisis based on the timely, adequate and consistent information attained from the main information channels. Satisfaction of security needs refers to the effective sense attained from effectively coping with the crisis and safe subjective experience, which is formed on the basis of the full understanding of the development of the public crisis and the assessment of the social assistance system and the public emergency capability.

2.2 Overview of Domestic and Foreign Research

Foreign research has gone through two stages: the first stage is the public sense of security in the perspective of "crisis management". "Bounded Rationality" of Simon [7] thinks that due to the limitation of the individual in memory, thinking, computing power and other aspects, individual rationality is just bounded rationality in constrained conditions. The second one is to introduce the theory of "risk perception" with the public sense of security examined by "crisis communication" model. Industrial pollution happening in 1984 in Union Carbide Corporation triggered panics due to slow release of information, which then urged the scholars to reflect. Covello finds that the risk assessment by the authorities is very different from the risk perception by the public and also the public show great distrust in the authority of risk management, which weakened their sense of security and even lead to fear [1]. Sun Ding and Matthew, the American scholars, thinks that the public sense of security is "the degree of anxiety and concern of those who are becoming victims", so the potential victims is selected as the objects in their research. However, they also agree that the public sense of security is a psychological phenomenon, which is characterized by anxiety or concern. Lois Mok, an American scholar, put more focus on the power of control in the sense of security and think it is the reflection on the normal order and the weakened social control from of the public [9].

In recent years, we have conducted deeper research on the public sense of security in the crisis, and the theoretical circle has shown a basic trend of multidisciplinary research and innovation. In terms of the social and psychological impact from critical events, Zhang Yan, in her study of the critical events, divided the psychological impact on the victims from critical events into three stages: the first stage is the formation of risk experience; the second is psychological cognition of risk; the third one is the implementation of mental decision and the spread of psychological influences [10]. The most important empirical investigations of the public sense of security was the sample survey conducted by Institute of public sense of security in the Ministry of public sense of security in 1988 in 15 provinces and urban areas with its results published in the book–"Do you feel safe?". The group defined the public sense of security as: the subjective feelings and evaluation of the public for the social security are the comprehensive psychological reflection of the infringement that the personal rights, property rights and other lawful rights and interests of citizens have suffered or will suffer and the protection they expect in a certain period of the social life [4]. Based on the theory of prospect, Sun Doyong constructed the research model to study the personal perception of fears in the crisis from three aspects: the features of the event itself, the characteristics of the individual and the social factors [8].

3 The Hypothesis in the Research

3.1 Build a Model of the Influencing Factors of Public Sense of Security

The critical events, the countermeasures of the government and the media, the emergency response capacity of the individual and the psychological behavior of the group in the natural disasters are described as four main factors that influence the public sense of security in "Research on the Influencing Factors of public sense of security and Countermeasures in Major Emergencies–An Empirical Analysis Based on the public sense of security of Ya'an Earthquake on April 20". Based on it, we assume that, in the disaster of security incidents, the critical events, the countermeasures of the government and the media, the emergency response capacity of the individual and the psychological behavior of the group are still the most important factors that influence the public sense of security.

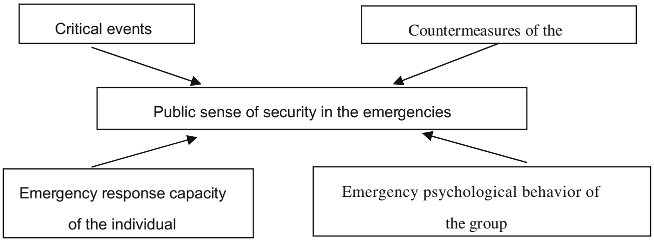


Fig. 1. Model of the influencing factors the public sense of security in the emergent events

Thus, we construct a conceptual model that affects public sense of security in the disaster of security incidents, which is shown as in Fig. 1.

3.2 The Hypothesis in the Research

Caplan and Edward assumed that each person is basically in a state of dynamic balance with the environment. When a person is facing a huge disaster in his life, he will have a high degree of tension, anxiety, pessimism and disappointment and a series of psychological problems if the disaster goes beyond his psychological endurance [2]. Thus, we assume that:

Hypothesis 1: In an emergency, the severity of the crisis, the extent of the impact and the number of casualties are negatively related to the level of the public sense of security. The more serious the crisis are, the more insecure the public will feel.

"The media dependence theory" put forward by Melvin Devler, a prominent American communication expert, points out that people are often anxious to get to know the truth through the government and the media, and the dependence will be enhanced obviously when the society is undergoing major changes but the situation is not clear. Therefore, the timely warning and the objective and fair reports from the government and the media is helpful to reassure the public and prevent further deterioration of the situation [6].

Hypothesis 2: In the emergencies, the countermeasures of government and the media to the emergencies is positively related to the public sense of security. The more timely and effective the countermeasures of the government and the media are, the safer the public will feel.

Robert Hiss put forward the FPC model and thought that the individual's ability and familiarity constituted the core variables of effective response to the crisis. The stronger the individual' ability and the higher the familiarity with the crisis are, the better the effect of dealing with the crisis will be [5].

Hypothesis 3: In the emergencies, the individual's knowledge of crisis, risk preferences, psychological quality are positively correlated to the public sense of security. The stronger the individual' emergency capacity are, the safer he will feel.

According to the theory of “Herd Behavior” which has been widely verified, individual in the crisis, often affected by others’ behavioral strategies, will often adopt the same behavioral strategy. That is to say, instead of being based on their own information, their behavior choice is the imitation of others or over-reliance on the public opinion, which tends to lead to the herd behavior. Thus, we assume that:

Hypothesis 4: In the emergencies, the psychology and behavior of the surrounding people are positively related to the public sense of security. The more stable the emergency psychological behavior of the group is, the safer the public will feel [8].

4 Research Methods and Tools

4.1 Scale

The paper collects data from the victims suffering from in the “explosion that happened in October 12 in Binhai New Area in Tianjin” by questionnaire survey which is divided into three parts. The first part is about the measurement of public sense of security and it makes the subjects score their sense of security; the second part is the measurement of public sense of security in emergencies; the third part is involved in the investigation of the five factors that influence hypothesis to finally form the questionnaire for formal study which is consisted of 25 items by the analysis of the factor and project.

4.2 Descriptive Statistical Analysis

A total of 330 questionnaires were distributed and 320 of them were collected with a recovery rate at 96.9%. After removing the unqualified questionnaire, 305 valid questionnaire were collected with the effective rate at 95.3%.

4.3 Examination of the Internal Consistency

According to the third part of the questionnaire, SPSS was used to analyze the internal consistency of the samples, which is shown as in Table 1. The results show that the alpha coefficient is 0.839, which proves that the designed questionnaires have good internal consistency and also indicate that the design of the questionnaire is scientific and reasonable.

Table 1. Analysis results of the internal consistency

Alpha of Cronbach	Alpha of Cronbach that is based on standardized project	Number of the projects
0.839	0.837	20

5 Results

5.1 Analysis of the Impact of Gender on the Public Sense of Security

In this questionnaire, we conducted survey on the demographic characteristics of respondents from the perspective of gender.

Using Independent-samples T test of the SPSS, we analyzed the impact of gender on sense of security with its results shown in Table 2. In the first part, we tested Levene' variance homogeneity with $F = 0.041$, $p = 0.084 > 0.05$, which meant that the variance was homogeneous while in the second part, the tested $t = 2.585$, $p = 0.010 < 0.05$, symbolized that the mean difference of the two samples is significant.

Table 2. Results of T test about gender

	Levene test in variance equation			T test in mean equation					
	F	Sig.	T	df	Sig. (in two sides)	Difference between the means	Standard margin of error	Min.	Max.
Index of the sense of security construct the equal variances	0.041	0.841	2.585	314	0.01	0.609	0.236	0.146	1.073
Construct the unequal variances	0	0	2.569	298,915	0.011	0.609	0.237	0.143	1.076

The results of T test show that the public sense of security is significantly affected by the gender after the explosion and the sense of security of men is distinctly higher than that of women, which may be explained by the environment where men and women grew up or their physical differences and also shows that in the post-disaster relief, we need to put more focus on women to rescue them.

5.2 Exploratory Factor Analysis

SPSS was used to carry out exploratory factor analysis on the samples with its results shown in Fig. 3. v13 to v32 are the corresponding items of the four influencing factors in the questionnaire. Among them, v13 to v15 correspond to critical events, v16 to v23 to the countermeasures of the government and the media, v24 to v27 to the emergency response capabilities of the individual, and v28 to v32 to the emergency psychological behavior of the group.

The results of exploratory factor analysis showed that the influencing factors mainly concentrate on four aspects, which was consistent with the hypothesis of the paper. Because of the dispersion of the concentration of the emergency psychological ability of the group in exploratory analysis, after analyzing and interpreting the corresponding items, we deleted two of them to conduct the exploratory factor analysis once again, which showed that the polymerizing effect of the four factors was good.

As shown in Table 3

Table 3. Results of exploratory factor analysis

	Elements			
	1	2	3	4
v13	-	-	-	0.852
v14	-	-	-	0.777
v15	-	-	-	0.619
v16	0.778	-	-	-
v17	0.822	-	-	-
v18	0.775	-	-	-
v19	0.747	-	-	-
v20	0.674	-	-	-
v21	0.725	-	-	-
v22	0.687	-	-	-
v24	-	-	0.646	-
v25	-	-	0.479	-
v26	-	-	0.630	-
v27	-	-	0.781	-
v28	-	0.641	-	-
v29	-	0.508	-	-
v30	-	0.680	-	-
v32	-	0.730	-	-

Attainment Method Element analysis of the subject. Pivoting method Formal maximum variation method with Kaiser.

5.3 Examination of the Model of the Influencing Factors of the Public Sense of Security

Based on the above four factors, this paper uses Amos to test the structural equation model, the indicators of which all meet expectation, as shown in Fig. 2. It can be seen that the four factors have a better explanation towards their impact on the public sense of security. Among them, only the impact of the critical events on the public sense of security is negative, indicating that they are negatively correlated, while the other three factors have positive impacts. Compared with the five hypotheses originally designed, the trust in the government and the countermeasures of the media are combined as one influencing factor. At the same time, according to the value of how these factors influence the public sense of security, we can see that the emergency psychological behavior of the group has the greatest impact, followed by the emergency capacity of the individual, the countermeasures of the government and the media and the severity of the critical events.

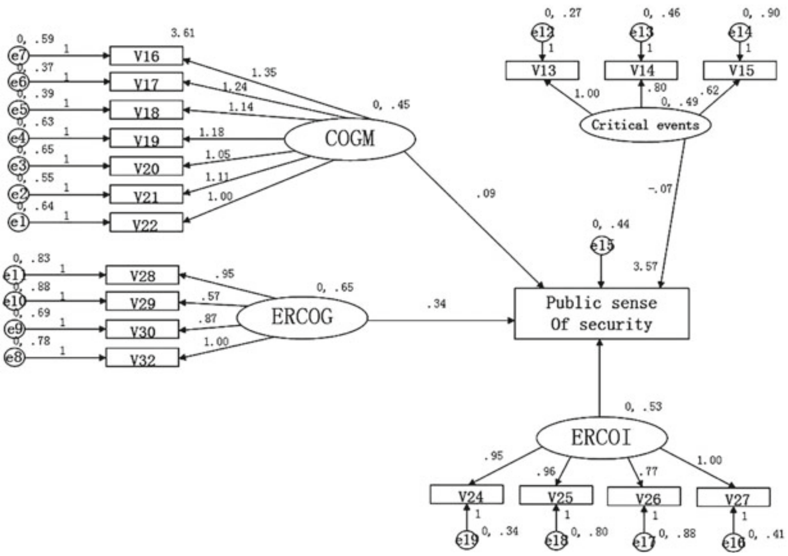


Fig. 2. Model of the influencing factors of the public sense of security. (ERCOG-Emergency response capacity of the group, COGM-Countermeasures of the government and the media, ERCOI-Emergency response capacity of the individual)

5.4 Examination on the Exploratory Factor Analysis of the Structural Dimension of the Public Sense of Security

We divided the structural dimension of the public sense of security, examined by SPSS, into sense of belonging, certain sense of control and sense of security with its results shown in Table 4: We can see that 10 items of the questionnaire are better aggregated into three categories, the coefficient of which statistically meet the requirements. We can see that V3, V4, V5, V6 come together as the first dimension while V7, V8, V9 as the second and V10, V11 and V12 as the third, which is consistent with the idea of how questionnaire is designed, but the structure of the questionnaire is more clear. According to the design of the questionnaire, we reduced the first dimension to a sense of belonging and the second dimension to a sense of security and the third to certain sense of control.

Table 4. Exploratory factor analysis of the structure dimension of the public sense of security

	Rotation matrix of the components		
	1	2	3
v3	0.708	-	-
v4	0.841	-	-
v5	0.777	-	-
v6	0.557	-	-
v7	-	0.608	-
v8	-	0.860	-
v9	-	0.815	-
v10	-	-	0.752
v11	-	-	0.730
v12	-	-	0.800

5.5 How the Influencing Factors of the Public Sense of Security Affect the Sense of Security

On the basis of determining the model of the influencing factors of the public sense of security and how each factor affects the public sense of security, we explored the mechanism of how the influencing factors affect the structural factors of the sense of security.

(1) How the influencing factors of the public sense of security affect the sense of belonging

First of all, we tested the structural equation model based on how the influencing factors of the public sense of security affect its first structural factors–"sense of belonging" with its results shown in Fig. 3. The results that may be affected by the characteristics of the security incidents showed that the emergency psychological behavior of the group had the strongest effect on the sense of belonging of the individual while the emergency response capacity of the individual comes second, followed by the countermeasures of the government and the media. When accidents and disasters that human error should be to blame for occurs, the credibility of the government and the media in the public mind will somewhat decline. The public will be more dependent on how the surrounding people deal with disasters, the rational behavior and effective countermeasures of whom will give the public a "sense of belonging." With the development of the newly emerging net media, the knowledge of how to effectively deal with the disaster has been more and more universal, at the same time the overall emergency response capacity of the group is also improved, which is consistent with the reality.

(2) How the influencing factors of the public sense of security affect the security needs

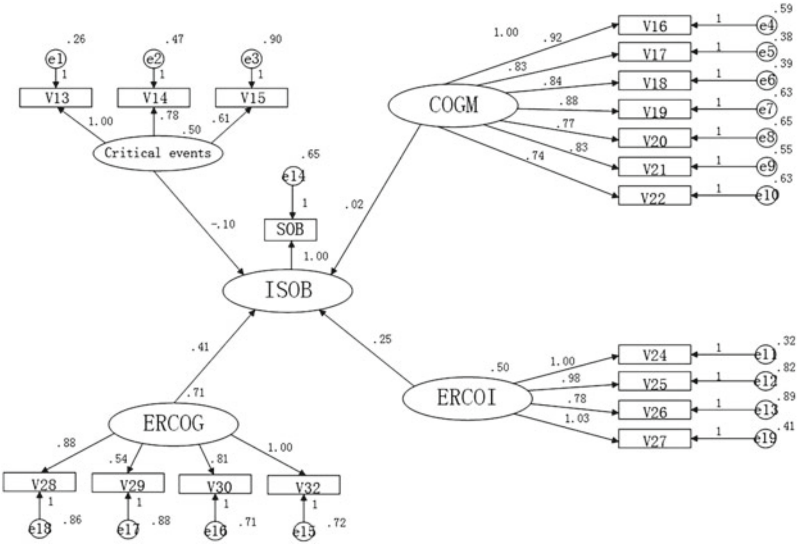


Fig. 3. Model of how the influencing factors of the public sense of security affect the sense of belonging. (ERCOG-Emergency response capacity of the group, COGM-Countermeasures of the government and the media, SOB-Sense of belonging, ISOB-Individual's sense of belonging, ERCOI-Emergency response capacity of the individual)

First of all, we tested the structural equation model based on how the influencing factors of the public sense of security affect its second structural factors–"security needs" with its results shown in Fig. 4. The results showed that the emergency response capacity of the individual was the most important factor that affects the sense of security of life, followed by the emergency response capability of the group and the countermeasures of the government and the media. Although personal characteristics, knowledge of emergency response and reactions on the spot are vastly different in the individuals, the security needs, for us, is the most essential experience of our safe lives and our emergency response capacity can directly affect our sense of security. In spite of the fact that the personal characteristics are difficult to control, it is possible to maximize the rationality by improving the personal crisis knowledge and his emergency response capacity in the face of major emergencies and minimize the casualty through effective emergency response. Once the individual has the clear understanding of that he should take immediate effective emergency measures, he will maximize the index of his sense of security. At the same time, the results also showed that in the crisis the emergency psychological behavior of the surrounding people will have great effect on the security experience of the individual.

(3) How the influencing factors of the public sense of security affect the certain sense of control

First of all, we test the structural equation model based on how the influencing factors of the public sense of security affect its third structural factors–the

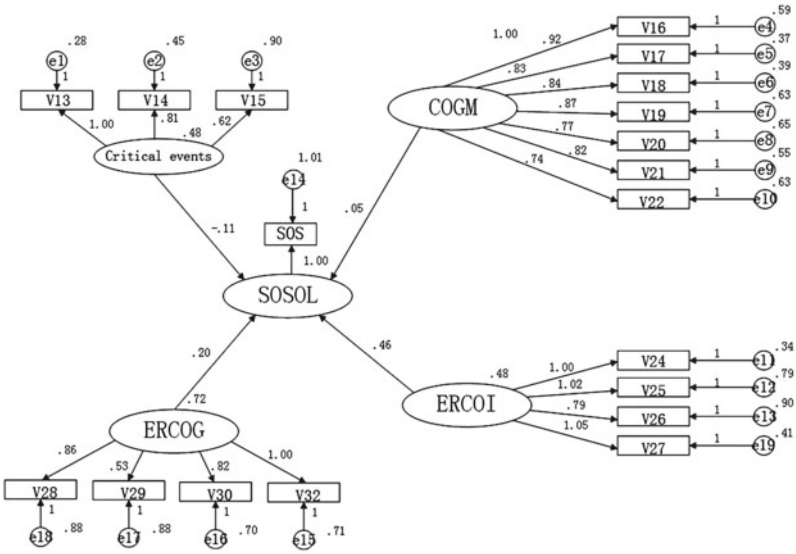


Fig. 4. Model of how the influencing factors of the public sense of security affect the security needs. (ERCOG-Emergency response capacity of the group, COGM-Countermeasures of the government and the media, SOS-Sense of security, SOSOL-Sense of security of life, ERCOI-Emergency response capacity of the individual)

"certain sense of control" with its results shown in Fig. 5. The results showed that the emergency psychological behavior of the group is the most important factor which can affect the certain sense of control of the public with the countermeasures of the government and the media coming second and the emergency response capacity of the individual third. The emergency behavior of the group can offer the demonstration and guidance to personal behavior, therefore, in the security incidents, their rational and effective emergency behavior can enhance the sense of control of the individual. In this mutual action model, the countermeasures of the government and the media also have greater influence, because it is a key to obtain the adequate, accurate and timely information for enhancing the sense of control. In China, our government and media have the ability to grasp the information of the emergencies and delivery them to the public adequately, timely and accurately to help themata in the whole information and improve their certain sense of control.

6 Countermeasures

In the major emergencies, the biggest problem that the public face is the loss of their sense of security, which, as a concept of subjective cognition, is to a large extent affected by self psychology. Thus, we can by certain means and ways affect our sense of security so as to enhance the effectiveness of disaster aid.

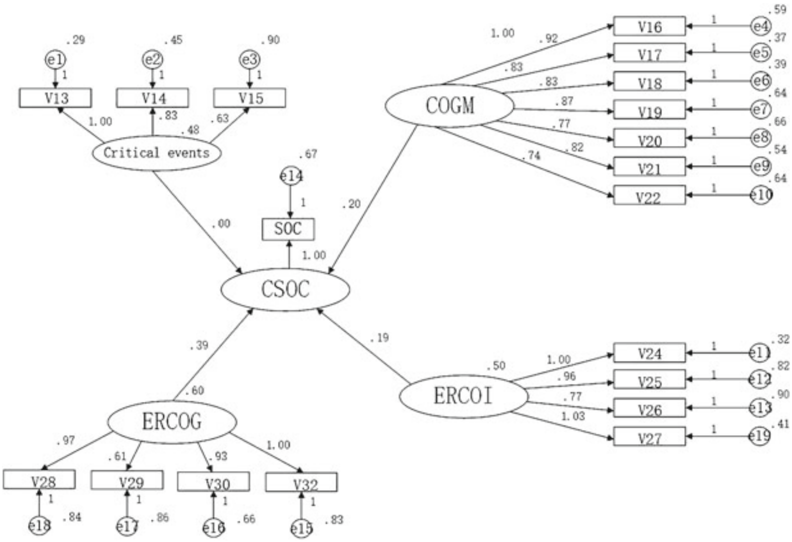


Fig. 5. Model of how the influencing factors of the public sense of security affect the certain sense of control. (ERCOG-Emergency response capacity of the group, COGM-Countermeasures of the government and the media, SOC-Sense of control, CSOC-Certain sense of control, ERCOI-Emergency response capacity of the individual)

When the crisis happens, it is a starting point of emergency psychological governance to create the public sense of security and the measurement of its influencing factors and how it is affected provide the evidence and direction for us to create a countermeasure system of social emergency psychological governance, which will be introduced in terms of the following three aspects.

6.1 Enhance the Countermeasures of the Emergency Response Capacity of the Group

In the measurement of the influencing factors of the public sense of security, the emergency response capacity of the group has the greatest impact. Personal behavior, influenced both by the external factors and group action, is easy to imitate the group action. In critical events, whether the emergency behavior of the group is effective or not will directly affect the sense of security of the individual. Therefore, it is necessary to enhance emergency response capacity of the group by regularly relevant knowledge and disaster drills.

6.2 Improve the Countermeasures of Certain Sense of the Crisis

In light of the countermeasures of the government and media as the key influencing factors of the public sense of security, unscientific reports on critical events from medias and the delay of the emergency measures from the government will lead to a serious decrease of the public sense of security and also fuel their fears. In the case of asymmetric information, government and the media should play their role of "delivering information timely" to enhance the sense of certainty of the public. The emergency measures taken by the government departments should be combined with the information distributed by the media and the view of the media to keep the information timely, adequate and accurate so that they can enhance the certain sense of crisis of the public.

6.3 Put More Importance to the Countermeasure of the Sense of Control of the Situation

This study also validate the assumptions that the nature and severity and other aspects of the critical events have impact on the emergency psychology of the public. Public fear in the face of crisis is an important reason for the loss of the sense of security. A sense of control of the situation should be created to help the public know how to deal with the crisis. They have become the important topics of emergency psychological governance to strengthen crisis education of the whole society so that people in the face of crisis can keep rational and orderly in the risk society. We should create a four-in-one crisis education system of the government, media, social organization and the public and design diagnostic measures, publicity measures and mentoring approach of the effective emergency measures in all kinds of situations for the public.

Acknowledgements. This paper is supported by the National Philosophy and Social Science Fund Youth Project (14CZZ017), a Study on the Influencing Factors and Construction Mechanism of Public Security in the Background of Major Incident, 2014/06–2016/12 2 and the Sichuan Province Twelfth Five-year Plan Philosophy Social Science Planning Fund Youth Project (SC11C002), Research on the Mobilization Mechanism of Internet Virtual Community Organization and the Countermeasures of Government. 2010/06–2015/12.

References

1. Covello VT, Winterfeldt DV, Slovic P (1987) Communicating scientific information about health and environmental risks: problems and opportunities from a social and behavioral perspective. In: Covello VT, Lave LB, Moghissi A, Uppuluri VRR (eds) Uncertainties in risk assessment and risk management, pp 221–239
2. Heath RL, Nathan K (1990) Public relations' role in risk communication: information, rhetoric and power. Public Relat Q 68(3):15
3. Lin Y (2007) Public sense of security and the construction of index system. Soc Sci 7:67–68
4. Ren Y, Wei J (2015) Analysis on the influencing factors of public concern in public crisis. Stat Decis 1:67–70
5. Robert M (2004) Heath, crisis management. Zhongxin Press, Beijing
6. Rocky B, Zheng Z, Wang B (2004) From "dependence on the media system" to "communication infrastructure"-looking back to the development of "media system dependency theory" and its new concept. Int Journalism 2:9–12

7. Simon HA (1978) On how to decide what to do. Bell J Econ 9(2):494–507
8. Yang J, Yang M (2016) Research on the influencing factors of public sense of security and countermeasures in major emergencies-an empirical analysis based on the public sense of security of ya'an earthquake on April 20. Exploration 1:172–179
9. Yang M (2015) Study on the influencing factors of public sense of security in major natural disasters. Master's thesis, University of Electronic Science and Technology of China
10. Zhang Y (2011) Risk attitude, risk perception and trust: an analysis framework of government information supplying mechanism in crisis situation based on prospect theory. J Huazhong Univ Sci Technol 25(1):53–59

Effect Assessment of the Free Tissue Flap Colorimetric Card in the Postoperative Flap Management for Tumors of Oral and Maxilla-Facial Region

Xiongtao Yang, Xueying Wang, and Qing Yang(✉)

Department of Head and Neck Surgical Oncology,
Sichuan Cancer Hospital, Chengdu 610041, People's Republic of China
57771050@qq.com

Abstract. The objective of this paper is to investigate the effect of free tissue flap colorimetric card in the postoperative flap management for tumors of oral and maxilla-facial region. Through observing 42 patients who used free tissue flap colorimetric card after operation, it was found that twelve patients had vascular disease, three had insufficient arterial infusion, and nine had venous return obstruction. Among them, one patient who had insufficient arterial infusion and two had venous return obstruction received re-anastomosis of vascular surgery and hemostasis surgery, and then all patients recovered well. The paper conclude that the free tissue flap colorimetric card is a simple, timely, accurate and effective tool for the observation of flaps, and it is suitable for clinical application.

Keywords: Free tissue flap colorimetric card · Tumors of oral and maxilla-facial region · Skin flap care · Crisis of blood-circulation

1 Introduction

Radical surgery for malignant tumors of oral and maxilla-facial region involves a wide range of injury, resulting in soft and hard tissue defects, seriously affecting the appearance and function of patients. Using their own tissue flap free graft to the mouth for large area of tissue defects repair has become a common method to improve the quality of life of patients [11]. It is the most important care content after free skin flap transplantation surgery to timely detect and actively deal with crisis of blood-circulation, as well as to promote the survival of the flap. In order to improve the success rate of free tissue flap, reduce the occurrence of crisis of blood-circulation and improve the success rate of the crisis tissue flap rescue, the branch designed the colorimetric card for free tissue flap in 2011 and applied the national utility model patent certificate (Certificate No. 3767860). Whats more, it has been used in has been in clinical application for more than

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_140

3 years, and the reflecting is very good. In this paper, the effect of free tissue flap colorimetric card in the postoperative flap management for tumors of oral and maxilla-facial region was concluded.

2 Materials and Methods

2.1 Normal Information

From September 2013 to November 2016, 42 cases of free flap repair were performed in our department, including 35 males and 7 females with an average age of 46.2 years (range, 17–74 years), including 17 cases of tongue cancer, 9 cases of buccal cancer, 7 cases of oral bottom cancer, 5 cases of gum cancer, 4 cases of soft palate cancer.

The information of surgery: there were 27 forearm radial free skin flap in repairing of oral cavity defect, 6 anterolateral thigh flap in repairing of oral cavity defect, 4 free pectoralis major myocutaneous flap in repairing of oral defect, 3 fibula myocutaneous flap in repairing of oral and maxillofacial defect, 2 osseo myocutaneous iliac flap in repairing of oral and maxillofacial defect.

2.2 Observation and Nursing of Skin Flap

Free flap is mainly observed blood flow of the skin flap, such as the color of the flap, temperature, degree of swelling and Shiatsu back to the blood reaction so as not to crisis of blood-circulation. Observer should avoid flashlight and other light interference, while maintaining the correct posture, to prevent the flap tension is too large or distorted [12]. The flap was observed by free tissue flap colorimetric card immediately after the patient was returned to the ward, and the results were recorded on the free tissue flap monitor. The observation was recorded every 30 min within 24 h after surgery, every 2 h within 24–72 h after surgery, every 4 h within 72 h–7 d after surgery. Abnormal flaps were observed every 15 min to record, and observe the records as usual since removing crisis of blood-circulation.

2.3 The Design of Free Tissue Flap Colorimetric Card

There are 10 colorimetric card for free tissue flap grafting made of silicone, its four corners rounded. The color areas of the ten cards are selected according to Table 1 and are different from each other, like Fig. 1.

The connecting shaft 6 has at least three elongated cards 1 whose one end is provided with a connecting hole 2, and the other end is provided with a colored area 3, and all the cards 1 are mounted on the connecting shaft through the connecting hole 2 in a manner similar to the fan structure, which is rotatable about the connecting shaft 6. The side edge of the card is also provided with a graduation zone 4 (accurate to millimeters) adjacent to the color zone 3 to accommodate various sizes within the oral cavity of the patient. In addition, the card is also provided with a text description area 5, with the color area 3 (seen Table 1).

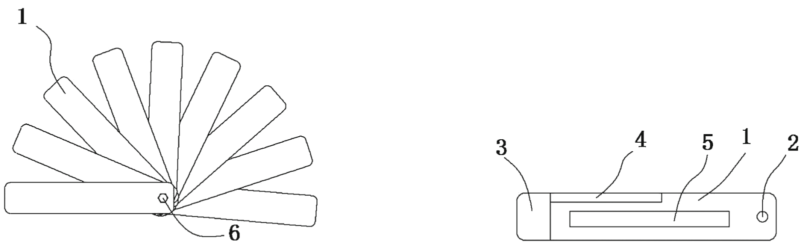


Fig. 1. The design of free tissue flap colorimetric card

Table 1. Description of color about free tissue flap colorimetric card

Parameter	Pale 1	Pale 2	Pale 3	Rosy 1	Rosy 2	Rosy 3	Violaceous 1	Violaceous 2	Violaceous 3	Atropurpureus
Tonal	20	20	20	20	9	9	234	235	225	225
Saturation	240	240	240	240	231	232	80	79	75	75
Brightness	226	216	204	195	166	152	169	146	126	45
Red R	225	225	205	255	252	252	205	188	172	63
Green G	240	230	217	207	136	113	155	122	96	33
Blue B	225	205	179	159	100	70	162	131	125	44

Note: the form of parameters are computer palette.

Pale1: arterial of flap flood inadequacy severally.
Pale2: arterial of flap flood inadequacy generally.
Pale3: arterial of flap flood inadequacy lightly.
Rosy1: the flaps grew well and the blood circulation was smooth relatively.
Rosy2: the flaps grew well and the blood circulation was smooth.
Rosy3: the flaps grew well and the blood circulation was very smooth.
Violaceous1: venous return is blocked lightly.
Violaceous2: venous return is blocked generally.
Violaceous2: venous return is blocked severally.
Atropurpureus: flaps is dead.

2.4 Detailed Description of Free Tissue Flap Colorimetric Card

The patient opened his mouth as fully as possible in order to expose the flaps in the irradiation conditions of natural light or fluorescent white light. At first, let the colorimetric card range by fan-out, compared and estimated flap color belongs to which card 1 color area 3. Then in order to improve the accuracy, the card 1 of the color region 3 is further inserted into the oral cavity of the patient, and the contrast is confirmed in close proximity. If colorimetric card was contaminated by the patient's oral secretions, blood, etc. in the usage of the card, it is necessary to be disinfected with alcohol or chlorine preparations after cleaning with water, and to prepare a set of colorimetric cards for every patients with specific infections. Colorimetric cards can also be set into a PVC gloves, discarded after every used. Generally, each card 1 is provided with a color region 3, and the number of the cards 1 can be determined according to the circumstances. When more accurate conditions are required, more cards 1 are required to make the whole structure have more color areas 3, in order to obtain more accurate flap discoloration.

2.5 Beneficial Effect of Free Tissue Flap Colorimetric Card

As the color areas were set in the card, the color of color area can be set by color in different conditions of flap, then the medical staff can objectively compare the transplanted flap color with the card on the color area in order to determine t whether there is a problem in the growth of the flap and to solve timely. This judgment of same standard reduces the judgment error caused by different subjective consciousness or lack of experience, which makes the flap observation simpler and more accurate.

3 Results

The 42 flaps were all live. Twelve patients had vascular disease, three arterial insufficiency, and nine venous return obstruction. Among them, 2 cases of mild arterial insufficiency maintain the correct position to release the pressure around the wound and the symptom relief after strengthening the local negative pressure drainage. 6 cases of mild venous return disorder were light by roasted lights. Among them, 2 cases relief after plus acupuncture flap bleeding, heparin sodium coated. 1 case of severe hypoperfusion of the arteries and 2 cases of severe venous return obstructed patients received re-anastomosis underwent vascular surgery and bleeding debridement surgery. 1 case of venous return to the middle and severe blocked, skin flap dead. All the patients recovered well after phlebotomy, and all the patients recovered well in phonation, chewing and swallowing (seen Table 2).

Table 2. The status about 42 cases of flaps after surgery

Situation of flap	# of cases	Vascular crisis time/h				Interventions measures			TFBBI	
		≤12	12–24	24–48	48–72	Lamp illumination	Needle bleeding of flap	Surgical debridement	≤6	≥6
Arterial insufficiency	3	1	1	1	0	0	0	1	1	2
Venous reflux disorder	9	0	2	6	1	9	2	2	2	7
Good growth of skin flap	30	0	0	0	0	0	0	0	0	0
Total	42	1	3	7	1	9	2	3	3	9

Note: TFBBI, Time of flap becoming better after intervention

4 Discuss

Large area of soft and hard tissue defects after the radical surgery for malignant tumors of oral and maxilla-facial region, often resulting in severe postoperative facial deformity and dysfunction. Although the survival rate of various free tissue flap has reached more than 90% [1,2,7,9], there are still a small number of patients whose tissue flap dead due to various reasons. The rate of crisis of blood-circulation is about to 3.5%–9.1% [8,10,14,15] according to domestic and foreign literature. Once the free tissue flap dead, it will give patients a disastrous blow, not only to extend the length of stay, increasing medical costs, more likely to make the mortality rate increased. The success of flap repair surgery, in addition to surgical techniques and equipment, but also thanks to good observation and nursing measures.

4.1 Routine Care of Flap After Surgery for Tumors of Oral and Maxilla-Facial Region

(1) Respiratory management

It is easy to lead to edema of Tongue and mouth bottom after surgery for tumors of oral and maxilla-facial region. The establishment and maintenance of airway patency is an important link, if necessary, preventable tracheostomy. There are a group of 19 patients with prophylactic tracheotomy. Strictly according to tracheotomy care routine, the use of saline and chymotrypsin atomization inhalation, 3 times a day, timely suction, no case of pulmonary infection occurred in patients. 1 case of buccal cancer free forearm flap transplantation, body fat, short neck, dyspnea occurred on postoperative day 3, and lift the respiratory distress after emergency tracheotomy.

(2) Position of care

Tissue flap transplantation area position nursing: Maintain the patient's head in the median or slightly biased side of surgery, non-distortions, effectively reducing the flap microvascular anastomosis tension. Raising bed 15°–30° after patients wake up is conducive to wound drainage and tissue valve venous return in order to reduce the surgical area and head and face edema. Attention to night patrol, and timely correct the patient is not correct position, to avoid the flap vascular pedicle compression, traction, discount.

Tissue flap for the position of care: Raising the donor limb elevation in order to facilitate venous return, and pay attention to finger blood supply and swelling. At the same time observe the wound bleeding situation, keep the dressing clean and Bandage intact. 1 patient of the group complaint the tissue flap for side limb numbness after 6 h, and the acral skin is pale by physical examination, skin temperature decreased, then immediately notify the physician to relieve decompression, symptom relief after rewrapping, to ensure the donor limb of the blood circulation.

(3) Ward management

Temperature of ward room is in 22–26° C, humidity 50% to 70%, and ward disinfect with ultraviolet every day 2 times. Attention to keep warm, to prevent crisis of blood-circulation due to vascular spasm by cold stimulation. Publicize non-smoking to family members and visiting staff, so that small artery spasm due to nicotine in tobacco, affecting the flap blood supply and induced arterial crisis.

(4) Medication care

It is important to attention to vital signs of observation, in particular changes in blood pressure which should be in 130/70 mmHg above, to ensure adequate blood volume. At the same time the usage of anticoagulation, anti-spasm, anti-infective drugs, routine use of low molecular weight heparin sodium, papaverine, dextran, Xiangdan should be as prescribed. In addition, observe the effect of medication and anticoagulant treatment side effects.

Strengthening nutrition promote incision growth. All patients underwent PG-SGA score, according to the results of the nutritionist to intervene, to provide patients with high-calorie, high protein and vitamin-rich diet plan, to prepare specialized nutrient solution if necessary. 42 patients retain the stomach tube to contact vacuum suction device, fasting on the day of surgery, to be drained liquid light green for nasal feeding which done 200 mL/time, interval 2 h 1 time. After the incision healed, the patient was instructed to practice swallowing. Nasogastric tube could be remove when oral feeding is without exception, instead of gold-type tube feeding (Using 50 mL syringe, a section of about 15–20 cm of the infusion tube, the tube set in the syringe, the injection of food when the skin placed in the mouth between the second and third molar). Diet from the liquid gradually transition to semi-liquid and solid food, avoiding cold, sharp, hard, with bone, barbed food.

(5) Nursing care of the irradiation of roasted lamps

This method can improve the local temperature of the flap, reduce arteriovenous vasospasm, as well as maintain microvascular smooth blood supply. With 40–60 W roasted lamp irradiation anastomosis, the use of distance 30–45 cm avoid baking lamp close-up flap, to prevent local temperature too high. There were nine cases of vascular crisis in this group, in the emergence of flap color darkening, flap tension increased to give roasted lights irradiation, with the use of vasodilator drugs in 2 cases after 6 h, 7 cases the flap color improved after 6 h.

(6) Nursing of wound drainage

Maintain the drainage area smooth, to prevent blood clots oppression arteriovenous effects of blood circulation and venous return. A one-time high vacuum drainage bottle contacting the end of drainage tube continued to attract, properly fixed drainage tube, regular squeeze, to avoid pipe distortion, compression, clogging, shedding, close observation of drainage fluid color, and measure the amount of drainage fluid and make a record 1 time/1 h within 24 h after surgery. If the drainage beyond 400 ml/d, bright red color, solidification, bleeding may be considered. If the amount is small, dark red, then alert subcutaneous congestion, and report to the doctor timely. If the drainage lower 10–15 ml/d, estuation can be considered after 48–72 h.

(7) Oral care

It was necessary to do oral management and to keep the flap area clean, if there were oral bleeding, more exudate after surgery. Routine oral care should be three times a day. Cotton balls should be along the line of blood vessels and pay attention to grasp the intensity when scrubbing. If necessary, to do oral cleaning. With 20 ml syringe aspiration of 3% hydrogen peroxide and saline as well as compound chlorine has been mixed, it was repeatedly washed, companying with a negative pressure. Note that suction should be on the contralateral, so as not to damage the flap and incision.

4.2 The Care of the Usage Free Tissue Flap Colorimetric Card in the Postoperative Flap Management for Tumors of Oral and Maxilla-Facial Region

After free flap anastomosis, a common complication is crisis of blood-circulation, divided into venous crisis and arterial crisis, usually occurring after 24–48 h [4]. The main monitoring indicators is the color, acupres sure and flap temperature difference [5]. The change of the flap color is the main basis for recognition and treatment. Early period of flap reconstruction, targeted nursing monitoring measures can help find vascular complications, providing first-hand information for clinical treatment. Weng [13] reported that flaps generally tolerate up to 6 h of ischemia without severe necrosis, 75% of the flaps would be necrosis after venous thrombosis continued 6 h and Flap necrosis rate would be 100% if beyond 8 h. Therefore, there some studies [6,13] pointed out that occurrence of flap ischemia and the timely monitoring of care is to determine the success of the key to the success of flap transplantation.

(1) Content of flap observation

Skin color changes and capillary reflow measurement, is the early minimum two indicators of reflection of blood circulation status of the most direct and rapid, subject to outside interference, and receiving repeat verification. Where both the skin color and capillary reflow abnormalities or more than two abnormal monitoring indicators, there is the crisis of blood-circulation. Venous crisis is usually manifested as red flap color, capillary response, and gradually become dark red, dark red, late black, acupuncturing out of dark red or black liquid when swollen hardened. Arterial crisis showed that the flap color is pale, lack of elasticity, acupuncturing without blood flow. However, the clinical description of the flap color and elastic temperature difference is often subjective, and the evaluation result is limited by the professional skill and nursing condition. First-line duty personnel for flap blood flow status judgment may be insufficient. If the flap color change with not knowing whether abnormal and the person do not have enough judgment experience, it may delay the flap vascular crisis rescue. We set up a free tissue flap management team, the department head deputy head appointed director and deputy director and head nurse as deputy head.

The backbone experts in the organization department access to a large number of domestic and foreign literature and review past patients pictures and case information, referred to the skin Flap vascular crisis indicators, combined with 20 years of clinical experience, to design the free tissue flap colorimetric card. In the flap color, degree of pale, ruddy, cyanotic were distinguished carefully to help clinicians to determine the extent of vascular crisis. At the same time, the observation of flap vision and abnormal changes in the scope of color flap provide accurate information for clinical care analysis, diagnosis and treatment according to the actual data. For 1 case of buccal cancer patients in the group after surgery 3 h, at 1:00 at night the flap color is pale 1 and skin temperature is slightly lower, as well as the examination body find no swelling of the neck, with plasma drainage tube patency and drainage of 15 ml dark red liquid. Doctor asked to record once every 15 min. The flap color transact pale 2 close to the pale 3 after 1 h and the neck was swollen with the drainage of the blood of the liquid drainage tube bright red liquid 50 ml. Then immediately report the surgeon, decisively send the patient to the operating room emergency line of neck incision debridement and vascular crisis exploration as well as small vessel anastomosis. After that, no longer found bleeding by observing 1 h, with the wound closed and a smooth period of 72 h risk period, the patients was back to the ward and at the latter part, voice, chewing, swallowing function of the patient recovered well.

(2) Observation field and time of flap

Defects in the postoperative flap management for tumors of oral and maxilla-facial region are generally located in the mouth. Especially the flap near the tongue observed inconveniently, so it need to increase the patient mouth opening, so that the full exposure of the flap. For 1 patient of the group within surgery 12 h, the wound edge bleed too much, the edge was dark purple and the area continue to expand to the center. It observed color change is not obvious within 2 s by sterile cotton swab compression capillary filling degree. And with 1ml syringe needle piercing the center of the flap, there were 0.5 cm dark red liquid outflow. Then we immediately notify the surgeon back to the operating room emergency treatment within 6 h after the flap color improved. There was another cases of oral bottom free radical repair, whose flap color was purple bark 1 and slightly faster response of the capillaries when back to the ward. Nurses found visible spots at mouth flap near the tongue 1/5 and increased speckle area with time in 13 h after surgery. When roasted lamp irradiation treatment is invalid, we immediately notify the doctor back to the operating room line exploration, and it found to be venous cramps combined embolism. Finally, the flap survived after the embolization vein re-anastomosis.

(3) The rule of crisis of blood-circulation

Flap crisis often occurs at night, which is linked with nocturnal high vagal tone, vascular intimal injury caused by hemodynamic changes, the role of nerve on the blood vessels and hormone secretion [3]. Nurses should strengthen the night patrol and attention to patient complaints as well as strict shift. There was 1 patients with left squamous cell carcinoma resection and left cervical lymph node dissection and thoracic duct ligation as well as common carotid artery epineurium and sublingual resection and free forearm flap excision repair and small vascular anastomosis and abdominal free skin graft repair complained neck pressure, pain, physical examination after 17 h. And physical examination only saw the flap suture and tongue bruising. But shift officers have not attracted great attention, so that the white skin flap color deepening, the area expanded to 1/2 or more when the shift after 6 h. It found negative pressure drainage tube shift, subcutaneous hemorrhage obvious, vascular anastomotic pressure and venous return blocked when sent to the operating room exploration. However, there was no obvious improvement in the flap condition at 6 h after operation, and the capillary reflex was not obvious. The elasticity also was poor and the skin temperature was low. Then the doctor surgery to trim the skin flap, to strengthen drainage, with local iodoformize gauze dressing and intravenous anti-infection as well as blood circulation and other treatment, so the flap survived.

5 Summary

Crisis of blood-circulation is one of the main complications of vascular anastomosis after vascular tissue transplantation. It is an important factor to timely find problems and take appropriate measures to ensure the survival of the flap. By using free tissue flap colorimetric card, it describe the flap color, capillary filling and the scope of flap color changes with the objective data as well as text, which makes the observation of free tissue flap more Simple and accurate. Therefore, it is worthy to promote in clinic.

Acknowledgements. This research was supported by the project from Sichuan Provincial Philosophy and Social Science Base and The key Research Base of Humanities and Social Sciences in Sichuan (XQ16C09).

References

1. Corbitt C, Skoracki RJ et al (2014) Free flap failure in head and neck reconstruction. Head Neck 36(10):1440–1445
2. Eckardt A, Meyer A et al (2007) Reconstruction of defects in the head and neck with free flaps: 20 years experience. Br J Oral Maxillofac Surg 45(1):11–15
3. Guo Q, Lu L, Xu M (2012) Observation of vascular crisis after transplantation of 12 cases of anterolateral flap. Chin J Nurs 47(3):215–217 (in Chinese)
4. Guo QY, Lu L et al (2012) Observation and nursing care of vascular crisis in 12 patients after anterolateral thigh flap transplantation. Chin J Nurs 47(3):215–216 (in Chinese)
5. Huang Z, Tao S, Hu L (2012) Nursing of 12 cases of complex tissue flap for large area soft tissue defect. Chin J Nurs 47(3):213–214 (in Chinese)
6. Jallali N, Ridha H, Butler P (2005) Postoperative monitoring of free flaps in UK plastic surgery units. Microsurgery 25(6):469–472
7. Kesting MR, Hölzle F et al (2011) Microsurgical reconstruction of the oral cavity with free flaps from the anterolateral thigh and the radial forearm: a comparison of perioperative data from 161 cases. Ann Surg Oncol 18(7):1988–1994

8. Lu M, Yang X, Wang Y (2010) Prevention and treatment of vascular crisis in free tissue flap of head and neck. J Oral Sci Res 26(2):243–245 (in Chinese)
9. Mao C, Yu G, Peng X (2002) Clinical analysis of 50 cases of free forearm flap for oral and maxillofacial. J Mod Stomatol. 16(2):138–140 (in Chinese)
10. Nakatsuka T, Harii K et al (2003) Analytic review of 2372 free flap transfers for head and neck reconstruction following cancer resection. J Reconstr Microsurg 19(19):363–369
11. Qiu W (2006) Recent progress in reconstruction of oral and maxillofacial surgery. Chin J Repar Reconstr Sur 20(4):339–343 (in Chinese)
12. Wang Q, Wang S et al (2010) Nursing care of 58 cases with microsurgical flaps for repairing soft tissue defects of legs complicated with. Chin J Nurs 45(8):702–704 (in Chinese)
13. Weng X, Wang J, Ye M (2010) Prevention and treatment of complications after maxillofacial defect repair. Chin J Stomatol Electron Ed 4(1):71–73 (in Chinese)
14. Ye WM, Zhang ZY et al (2004) clinical observation and treatment of failing flaps in oral and maxillofacial surgery: a report of 44 cases. Shanghai J Stomatol 13(2):91
15. Zhang C, Zhang Z et al (2005) [Repair and reconstruction of oral and maxillofacial defect-clinical analysis of 1973 cases]. Chin J Repar Reconstr Sur 19(10):773

The Impact of Intellectual Capital on High-Tech Enterprise Performance: Applied Study in China's Second-Board Market

Hongchang Mei(✉) and Kunling Wang

School of Administration, Chongqing Technology and Business University, Chongqing 400067, People's Republic of China
1426315339@qq.com

Abstract. This paper chooses high tech enterprises listed on China's Second-board Market as research objects. The aims of this thesis are to find out the impact of intellectual capital on the performance of high-tech enterprises. This study has carried out an empirical analysis with the method of multiple linear regression. The empirical findings indicate a trend that for high-tech industries on the GEM, the improvement of their performance efficiency depends more on the human capital compared to structural capital and physical capital. Both human capital and structural capital significantly influence the performance efficiency of enterprises in the high-tech industry, while physical capital practically doesn't impact it significantly. According to the empirical findings, this paper has put forward some advices on capital allocation for Chinese high-tech enterprises.

Keywords: Intellectual capital (IC) · Performance · High-tech enterprises

1 Introduction

In the industrial age, human capital is the driving force of economic growth, but in the age of knowledge economy, intellectual capital has become the power source of economic growth, which is also changing the rules of business and national competitiveness. Also, nowadays intangibles are almost universally considered as the main value drivers for companies [12].

China as a rapidly developing country, has strongly encouraged entrepreneurial activities especially in high-tech fields. Similarly, intellectual capital seems more important than any other factor in the fields of high-tech, that intellectual capital management have become the domain of the so-called Chief Knowledge Officer [1], there is a great need to study the impact of intellectual capital in Chinese high-tech enterprises. The result will not only aid in selection for managing intangibles, but also accelerate the pace of growing of high-tech companies.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_141

This research aims to investigate the impact of intellectual capital on the performance of selected enterprises in the high-tech sector on the Second-board market in China. The first section deals with the introduction of intellectual capital and related researches. This follows with research methodology and design. In the third section, the research empirical findings are presented. Finally, the conclusion outlines the limitations and implication of this study.

2 Literature Review

2.1 Intellectual Capital and Sub-Components of Intellectual Capital

The word Intellectual capital was first proposed by Senior as a synonym for human capital in 1836, which he believes is a combination of knowledge and skills of human beings. John Galbraith extended the concept of intellectual capital, noting that intellectual capital was not just knowledge or pure intellect, but also the process of making use of knowledge effectively. Guenther, Beyer [11] argued that intellectual capital represents the sum of all the intangible assets of the company. Also Intangible assets are defined as the company's non-material and non-financial resources, including technology, customer information, brand name, honor and corporate culture. The intellectual capital of an enterprise can improve the performance of an enterprise by relying on the dynamic interaction of IC components [4].

IC is also considered as the sum of following three categories: (1) Human capital, which refers to the people in an organization and describes their cumulative tacit knowledge skills. (2) Structural capital, which consists of supportive infrastructure, processes, and databases of corporation that enable human capital to function [8]. (3) Relational capital.

Another definition is that IC is derived insights about head value, future earnings capabilities, based on human capital, as well as organizational, structural and relational capital [5]. Some scholars hold the view that intellectual capital can be classified into human capital and structural capital, divides intellectual capital into human capital and structural capital, which is accepted by author of this article.

2.2 Relation Between IC and Firm Performance

Firer did an empirical research on the relationship between the three components of value added intellectual coefficient (VAIC) in South Africa. He concluded that, if the decision makers focus more on exploiting the physical capital reward, it would lead to the neglect of the use of human capital.

The human capital is very important to the enterprise performance, that companies should value it while managing their IC. Bontis [2] conducted a study on the efficiency of IC to enterprise value in service industry and other industry in Malaysian. The results show that human capital in various industries is an important role as structural capital and is positively correlated with the performance of enterprises. Fengetal [6] proposed that it was necessary to clarify the relationship between corporate performance and human capital with more attention to human capital.

They also have mentioned that performance of enterprises should be considered by combining the characteristics and degree of competition of specific industry. Zhu et al. [13] assumed that human capital is the key factor to determine enterprise performance.

2.3 IC Measurement Models

IC is difficult to measure as an intangible asset. So far there is no unified view on the measurement method for intellectual capital. The current measurement models are as follows.

(1) Difference between market value and book value, which assumes that the intellectual capital of enterprises could be the difference between their market value and book value.
(2) Tobin'sq, which refers to the ratio between a physical asset's market value and its replacement value. It was proposed in 1968 by James Tobin and William Brainard [3]. If the replacement cost of the company's assets is lower than the company's market value, the company's investment behaviors would have obtained excess profits, which are generated from IC [10].
(3) VAICAnte Public proposed Value Added Intellectual Coefficient (VAIC) in 1998 to measure the intellectual capital value of an enterprise. According to this method corporate capital consists of physical capital and intellectual capital, and corporate performance depends on the efficiency of value added by a firm from its physical capital and intellectual capital resources. This method assumed that the enterprise's intellectual value-added efficiency represented its performance.

Lentjušenkova and Lapina [7] proposed that, in a modern organization, its managerial approaches and investment facilities are changing along with the growing importance of IC and the changes of its components, thereby changing the managerial approaches and focusing on value creation.

The paper chooses the VAIC method to analyze impacts of intellectual capital on the performance of high-tech enterprises in China. In this sector, business in the following specific industries are preferred, like high-growth, high-tech, high value-added and new economy, new services, new agriculture, new energy and new materials, that possess high growth potential with high risk [9]. In these industries, the value and function of intellectual capital seems have exceeded financial capital.

3 Research Methodology and Design

This study is a longitudinal research in a time-span 2013–2015 and refers to China. Based on the developed VAIC model by Firer, this paper analyzes the correlation and regression between value-added coefficients of IC components and high-tech enterprise performance. And then explains the relationship between intellectual capital components and enterprise performance based on the analysis results. In this model, it is assumed that value creation of enterprise mainly comes from physical capital resources and intellectual capital resources.

3.1 Research Hypotheses

Economic growth mainly depends on physical capital and intellectual capital. For the composition of intellectual capital in this paper, it is assumed that IC includes human capital and structural capital. Human capital is the sum of knowledge, skills, experience of people in the corporate. IC can be performed in the products and services, knowledge level and education level of corporate members. This paper uses the index of human capital value-added coefficient to represent its intensity. Structural capital is an organizational asset that is embedded in organizational structure, management system standard and corporate culture. In this paper, the structural capital intensity will be represented by its value-added coefficient.

Hypothesis 1: The intellectual capital and physical capital of an enterprise cause the performance of an enterprise.
Hypothesis 2: Intellectual capital consists of human capital and structural capital.
Hypothesis 3: For high-tech enterprises, the impact of intellectual capital on corporate performance is significant.

3.2 Sample Selection and Data Sources

The sample of this article comes from the financial statements of 80 companies listed on the GEM from 2013 to 2015. Research samples in this paper are all certified as high-tech by related government agency and listed on China's second-board market in Shenzhen stock exchange from 2013 to 2015. In order to ensure the validity of the data, the criteria for screening the original samples are as follows:

(1) Selecting the samples of the listed GEM companies, which are in continuous operation during this period to ensure the necessary continuity of the sample set.
(2) Financial services listed companies are eliminated because of their large total share capital, which may cause deviation of the regression results.
(3) Excluding listed companies of which annual financial statement data are incomplete, or not continuously operating in the time-span 2013–2015.
(4) Elimination of abnormal operating listed companies (e.g. Debt ratio surpassing 100%) and companies with deteriorating financial condition to ensure the validity of the sample data.

(5) Listed companies who have changed their main business because of mergers and acquisitions or other acts will not be included in the sample set. The final sample set includes 80 high-tech enterprises. This article used EXCEL and EVIEWS9.0 in the processing of sample data.

3.3 Research Model

Enterprise's ability to creating value by using physical capital and intellectual capital is called "intellectual ability", which is expressed by VAIC model. This model consists of capital employed efficiency (CEE), human capital efficiency (HCE) and structural capital efficiency (SCE). Relation between them is that, $\mathrm{VAIC}_i = \mathrm{CEE}_i + \mathrm{HCE}_i + \mathrm{SCE}_i$.

The main purpose of this article is to study the impact of intellectual capital on the performance of high-tech enterprises. By establishing the multiple linear regression model and doing a statistical hypothesis test, we obtained the estimation for the regression coefficient. Based on all the above conditions, we will interpret the impact of IC on performance of Chinese high-tech Enterprises.

The basic model used to estimate the binomial logistic regression models is:

$$\mathrm{ROA} = \alpha + \beta_1 \mathrm{CEE} + \beta_2 \mathrm{HCE} + \beta_3 \mathrm{SCE} + \beta_4 \mathrm{Leverage} + \beta_5 \mathrm{LNSIZE} + \varepsilon.$$

4 Research and Discussion

4.1 Variable Selection

(1) Selection of dependent variable
Interpretations of selected variables in this paper are as follows:

Table 1. Dependent variable

	Variable symbol	Variable explanation	Variable expression
Dependent Variable	ROA	Return on Assets	Net Income after Taxes/Assets

ROA shows the percentage of how profitable a company's assets are in generating revenue. In the following analysis we use ROA to express the performance of a company (Table 1).

(2) Selection of independent variables
(See Table 2)

① CE is for the book value of total assets of an enterprise.
② HC represents human capital of an enterprise which is measured by total wage cost of an enterprise. The total wage cost in this article is indicated by cash paid to employees as salary and other payments in cash flow statement of a listed company.

Table 2. Independent variables

	Variables symbol	Variables explanation	Variable expression
Independent variables	CEE	Capital Employed Efficiency	CEE = VA/CE
	HCE	Human Capital Efficiency	HCE = VA/HC
	SCE	Structural Capital Efficiency	SCE = SC/VA

③ VA expresses the value added of enterprise, which is a sum of net income, income tax and total wage cost.
④ Income taxes cost is used in place of income tax in this thesis.
⑤ With income tax cost as a approximation of income tax.

(3) Selection of control variables
(See Table 3)

Table 3. Control variables.

	Variables symbol	Variables explanation	Variable expression
Control variables	LNSIZE	Logarithm of Main Business Income	SIZE = LN (Total assets)
	Leverage	Debt ratio	Leverage = Total debt/Total assets

① The size of the assets impacts on the value of enterprises. IT is usually expressed by the total assets of enterprises. In order to reflect the size of the assets of enterprises accurately, this paper selects the logarithm of main business income as a control variable.
② Debt Ratio is a financial ratio that indicates the percentage of a company's assets that are provided via debt. It can be used to measure the ability of a firm to operate. Therefore, we use the Debt Ratio together as the second control variable.

4.2 Model Construction

On the basis of previous theoretical analysis, this paper constructs a multiple regression model between physical capital efficiency, human capital efficiency, structural capital efficiency and firm performance. The model is set as follows:

$$\mathrm{ROA} = \alpha + \beta_1 \mathrm{CEE} + \beta_2 \mathrm{HCE} + \beta_3 \mathrm{SCE} + \beta_4 \mathrm{Leverage} + \beta_5 \mathrm{LNSIZE} + \varepsilon.$$

Here: α is for the constant, $\beta_1 - \beta_5$ are corresponding coefficients. ε represents the random disturbance.

Table 4. Descriptive statistics of variables.

Observations	2013		2014		2015	
	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev
ROA	0.061608	0.070561	0.065428	0.097572	0.068864	0.063246
LNSIZE	20.20134	0.855858	20.38727	0.888847	20.59145	0.8757
Leverage	0.27982	0.329841	0.224613	0.227157	0.304221	0.320626
CEE	0.174146	0.155064	0.213168	0.232301	0.196046	0.250991
HCE	2.095241	1.147357	2.430141	3.53092	2.484896	4.281109
SCE	0.309643	0.658739	0.41756	0.340202	0.343141	0.790112

4.3 Empirical Analysis

(1) Sample description statistics
We have described statistical relationships between variables in the model, that empirical results are as Table 4.

From results of descriptive statistics, we can see that the three-year average of ROA of high-tech enterprises on the China's second-board market are 0.061608, 0.065428 and 0.068864, which show that the performance of the high-tech enterprises listed on the China's second-board market is increasing year by year. The means of CEE in the sample period respectively are: 0.174146, 0.213168 and 0.196046, which means the change of physical capital is not very stable. But their standard deviations are: 0.155064, 0.232301 and 0.250991, which shows an increasing trend with time. It indicates that the difference between impacts of physical capital on each enterprise's performance is widening per year. The mean 2.095241, 2.430141 and 2.4484896 of HCE over sample period, which is showing an increasing trend, indicates that the influence of human capital on high-tech firms' performance listed on Chinese Second-board market is gradually increasing. The mean 0.309643, 0.417560 and 0.3343141 of SCE in these three years is showing an unstable trend without regularity.

We can also see from the above figure that mean of HCE is obviously greater than the mean of CEE and SCE during the period from 2013 to 2015, which tells that human capital of high-tech enterprises contribute more to the corporate performance than physical capital and structural capital.

(2) Partial correlation analysis for variables
In order to make the multiple regression analysis meaningful, we have analyzed the correlation between selected variables. The analysis results are as follows (Table 5):

It can be seen from the above analysis, that, there is a positive correlation between HCE, SCE, CEE and ROA. However, the correlation coefficient between ROA and CEE in 2014 is 0.0638, with the concomitant probability of T test 0.5712, which hasn't passed the confidence level of 1% test. This indicates that the impact of physical capital on firm performance is only a little in this year, The correlation coefficients of the other two years both correspond the 1% confidence

Table 5. Partial correlation analysis for selected variables.

			ROA	CEE	HCE	SCE	Leverage	LNSIZE
2013	ROA	Correlation	1	0.478	0.6902	0.4722	0.4403	0.6734
		Probability	-	0	0	0	0	0
	CEE	Correlation	0.478	1	0.75	0.4914	0.712	0.7435
		Probability	0	-	0	0	0	0
	HCE	Correlation	0.6902	0.75	1	0.634	0.5581	0.8792
		Probability	0	0	-	0	0	0
	SCE	Correlation	0.4722	0.4914	0.634	1	0.3012	0.4356
		Probability	0	0	0	-	0.0063	0
	Leverage	Correlation	0.4403	0.712	0.5581	0.3012	1	0.6431
		Probability	0	0	0	0.0063	-	0
	LNSIZE	Correlation	0.6734	0.7435	0.8792	0.4356	0.643135	1
		Probability	0	0	0	0	0	-
2014	ROA	Correlation	1	0.0638	0.4471	0.4426	0.342	0.5711
		Probability	-	0.5712	0	0	0.0018	0
	CEE	Correlation	0.0638	1	0.427	0.6259	0.6802	0.6705
		Probability	0.5712	-	0.0001	0	0	0
	HCE	Correlation	0.4471	0.427	1	0.6236	0.3954	0.5731
		Probability	0	0.0001	-	0	0.0003	0
	SCE	Correlation	0.4426	0.6259	0.6236	1	0.5557	0.7725
		Probability	0	0	0	-	0	0
	Leverage	Correlation	0.342	0.6802	0.3954	0.5557	1	0.6984
		Probability	0.0018	0	0.0003	0	-	0
	LNSIZE	Correlation	0.5711	0.6705	0.5731	0.7725	0.6984	1
		Probability	0	0	0	0	0	-
2015	ROA	Correlation	1	0.4335	0.4596	0.3742	0.4386	0.7515
		Probability	-	0.0001	0	0.0006	0	0
	CEE	Correlation	0.4335	1	0.418	0.3674	0.8234	0.6076
		Probability	0.0001	-	0.0001	0.0007	0	0
	HCE	Correlation	0.4596	0.418	1	0.3517	0.456	0.508
		Probability	0	0.0001	-	0.0013	0	0
	SCE	Correlation	0.3742	0.3674	0.3517	1	0.3422	0.3996
		Probability	0.0006	0.0007	0.0013	-	0.0018	0.0002
	Leverage	Correlation	0.4386	0.8234	0.456	0.3422	1	0.6803
		Probability	0	0	0	0.0018	-	0
	LNSIZE	Correlation	0.7515	0.6076	0.508	0.3996	0.6803	1
		Probability	0	0	0	0.0002	0	-

level of T test. At the same time, there are positive correlations between HCE, SCE and CEE, and all of them have passed T test with confidence level of 1%. This result tells that influencing factors of business performance are not independent of each other, but interacting with each other.

(3) Multiple regression

The regression results have shown that the results are not very ideal. Only the value of the intercept term in there gression equations passed the T test for linear regression (here 5%), with CEE, HCE, SCE coefficients did not. This result tells that the coefficients are not very credible. However, the F-test values 6.107650, 9.607723 and 5.317085 of each model have passed the F-test for linear regression (here 5%), which indicates explained variables in the model could be explained by explanatory variables (Table 6).

Table 6. Multiple regression model.

Time		2013		2014		2015	
Dependent variable		ROA		ROA		ROA	
		Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Constant		−0.6637	−3.7726 (0.0003)	−0.476358	−2.175358 (0.0328)	−0.645427	−3.986521 (0.0002)
Independent variables	CEE	−0.0602	−1.1093 (0.2709)	−0.237977	−5.567514 (0.0000)	0.04047	1.134204 (0.2604)
	HCE	0.0177	2.3107 (0.0236)	0.004414	1.623358 (0.1088)	0.001087	0.703479 (0.484)
	SCE	0.005	0.3683 (0.7137)	0.03222	1.119615 (0.2665)	0.008388	1.020737 (0.3107)
Control variables	Leverage	0.034184	1.427739 (0.1576)	0.077395	1.812786 (0.0739)	−0.025991	−0.927511 (0.3567)
	LNSIZE	0.034031	3.966973 (0.0002)	0.027024	2.563172 (0.0124)	0.034416	4.429544 (0.0000)
Adjusted R-squared		0.244296		0.352664		0.214598	
F-statistic		6.10765		9.607723		5.317085	
Durbin-Watson stat		1.626685		2.100722		1.85182	

Besides, we can see that the coefficients of CEE in 2013 and 2014 respectively are −0.0602 and −0.237977, with 0.040470 in 2015, which indicates that CEE has a negative effect on ROA. Coefficients of other explanatory variables to ROA are positive, which tells that human capital and structural capital drive performance of high-tech enterprises. DW 1.626685, 2.100722, 1.851820 in the time-span are all close to 2, indicating that the regression equation does not have strong autocorrelation.

5 Conclusion

5.1 Limitations

The results seem to be achieved, but there are some limitations of the research. The first concerns the fact that the sample set limited to listed high-tech companies, because their financial data are available, while financial data of non-listed high-tech companies are not. Accordingly, carrying out further research linking IC and non-listed high-tech enterprises could develop this research field.

Another limitation concerns that some indicators of basic data used substitutions due to the availability of fundamental data on each company's financial report. So it needs more in-depth, exhaustive analysis to be carried out. A further limitation is, this analysis considers only the time-span 2013–2015 without consideration of environmental differences.

5.2 Implications for High-Tech Enterprises in China

From the above empirical research, we find that human capital and structural capital have significant positive correlation with the performance of listed high-tech enterprises, while physical not significant. For the high-tech enterprises listed on China's GEM, human capital of intellectual capital has greater impact on the corporate performance than structural capital and physical capital does. Structural capital has relatively less effect compared to other human capital on high-tech firms' performance. However, high-tech enterprises still need to pay attention to structural capital, and strengthen enterprise management. The average value-added efficiency of intellectual capital is bigger than that of physical capital, which reflects that the dependence of enterprise performance on intellectual capital is increasing in high-tech field. The author puts forward based on this empirical research following suggestions:

(1) High-tech enterprises should devote more efforts to the development and management of the intellectual resources of employees, and give full play to the intelligence of them.
(2) The focus of high-tech enterprises should be transferred from physical capital to intellectual capital.
(3) Designing enterprise size according to IC of a company will help to operate structural capital and physical capital more reasonable, and thus enhance their efficiency of creating corporate value.

References

1. Bontis N (2001) CKO wanted-evangelical skills necessary: a review of the chief knowledge officer position. Knowl Process Manage 8(1):29–38
2. Bontis N, Keow W, Richardson S (2002) Intellectual capital and business performance in malaysian industries. J Intell Capital 1(1):85–100
3. Brainard WC, Tobin J (1968) Pitfalls in financial model building. Cowles Found Discuss Pap 58(2):99–122

4. Chen Y (2011) Formation and development of “new over-the-counter market”. China Secur Futures China 11:15–24 (in Chinese)
5. Edvinsson L (2013) IC 21: reflections from 21 years of ic practice and theory. J Intell Capital 14(1):163–172
6. Feng L (2007) An empirical analysis on the relationship between human capital and firm performance. Commun Finance Account 2:66–69
7. Lentjušenkova O, Lapina I (2016) The transformation of the organization’s intellectual capital: from resource to capital. J Intell Capital 17(4):610–631
8. Maddocks J, Beaney M (2002) See the invisible and intangible. Knowl Manage 3:16–17
9. Meiou L (2013) Research on China gem development. Master’s thesis, Fudan University, Shanghai (in Chinese)
10. M’Pherson PK, Pike S (2001) Accounting, empirical measurement and intellectual capital. J Intell Capital 2(3):246–260
11. Subramaniam M, Youndt MA (2005) The influence of intellectual capital on the types of innovative capabilities. Acad Manag J 48(3):450–463
12. Zambon S (2016) Ten years after: the past, the present and the future of scholarly investigation on intangibles and intellectual capital (IC). J Intell Capital 17(1):42–50
13. Zhu X (2003) Human capital and corporate performance. Guangxi Soc Sci 1:75–77 (in Chinese)

Site Selection of Public Fast Electric Vehicle Charging Station by Using an Intuitionistic Fuzzy Projection-Based Approach

Lin Zhong[1] and Zhibin Wu[2](✉)

[1] Uncertainty Decision-Making Laboratory, Sichuan University, Chengdu 610064, People's Republic of China
[2] Business School, Sichuan University, Chengdu 610064, People's Republic of China
zhibinwu@scu.edu.cn

Abstract. As an important solution to reduce the dependency on crude oil and minimize transportation-related carbon dioxide emissions along with other pollutants, electric vehicles (EVs) are broadly promoted by government. As the energy provider of EVs, electric vehicle charging stations (EVCSs) play important roles in electric vehicle popularization. The effective site selection of EVCS is in favour of promoting the willing of consumer to use EVs. In this paper, we apply multi-criteria group decision-making (MCGDM) method to site the public fast electric vehicle charging stations. To avoid the information lost in aggregation process, we use an intuitionistic fuzzy projection-based approach. The proposed method is applied into a numerical example, and the results show that the proposed MCGDM is feasible and efficient for EVCS site selection.

Keywords: Intuitionistic fuzzy set · Electric vehicle · Fast charging station · Multi-criteria decision making algorithm

1 Introduction

Electric vehicles (EVs) are considered as a significant solution to reduce the dependency on crude oil and minimize transportation-related carbon dioxide emissions along with other pollutants [7]. As the energy provider of EVs, electric vehicle charging station (EVCS) is the foundation of electric vehicle industry development [4]. Efficient, convenient and economic EVCS is in favour of promoting willingness of consumers and also the industry development. As the preliminary work of EVCSs constructions, the EVCS site selection is quite important in the whole life cycle, which has significant impacts on the service quality and operational efficiency of EVCS. Therefore, it is necessary to employ proper methods to determine the optimal EVCS site.

Li et al. [6] considered dynamic origin-destination trip satisfaction and public electric vehicle (EV) charging network expansion to develop a multi-period

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_142

refueling location model. Sheppard et al. [8] applied an agent-based simulation modeling platform to site EVCSs in Delhi. Antunes et al. [3] used a maximal coverage model to determine the optimal location, quantity and capacity for EVCSs. Although many models have been proposed to site EVCS, few papers have considered multi-criteria group decision making (MCGDM) techniques.

In classical decision-making problems, the preference value for each alternative expressed by each decision maker is precise [2]. Due to the fuzziness and uncertainty of decision-making problems and the inherent vagueness of human preferences, however, the best way for decision-makers to give their opinion is using natural language [11]. For that reason, linguistic variables [18], the value of which are expressed using words and sentences from natural language, have been widely used to indicate the preference values for each attribute. Aggregating linguistic information has generally been discussed under four main schemes [10].

The purpose of this paper is to apply the intuitionistic fuzzy projection-based approach to determine the optimal site for public fast electric vehicle charging station to promote the popularization of EVs in Chengdu. This framework differs from the traditional method as the weights of decision makers are undetermined and the complete process is conducted in an intuitionistic fuzzy environment. The remainder of this paper is organized as follows. In Sect. 2, we briefly review the basic concepts of intuitionistic fuzzy sets. The proposed fuzzy multi-criteria group decision making framework is presented in Sect. 3. Then, in Sect. 4, the proposed MCGDM is applied to determine the site for PFEVCSs. Finally, conclusions are drawn in Sect. 5.

2 Preliminaries

In this section, we briefly review the basic concepts of IFSs.

Definition 1. [17] Let X be a universe of discourse, then a fuzzy set A is defined as

$$A = \{\langle x, \mu_A(x)\rangle | x \in X, \mu_A(x) \in [0, 1]\}, \tag{1}$$

where $\mu_A(x)$ denotes the membership degree of the element x to the set A.

Definition 2. [1] Let X be a universe of discourse, then an intuitionistic fuzzy set A is shown as follows:

$$A = \{\langle x, \mu_A(x), \nu_{A(x)}\rangle | x \in X, \mu_A(x) \in [0, 1], \nu_A(x) \in [0, 1]\}, \tag{2}$$

where $\mu_A(x)$ and $\nu_A(x)$ respectively represent the membership degree and non membership degree of the element x to the set A with the condition $0 \leq \mu_A(x) + \nu_A(x) \leq 1$.

Definition 3. [1] For each intuitionistic fuzzy set A, if

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x), \forall x \in X, \tag{3}$$

then, $\pi_A(x)$ is called the indeterminacy degree or hesitant degree of x to the set A. The bigger the $\pi_A(x)$, the more the indeterminacy degree of knowledge about x [15]. Especially, when $\mu_A(x) + \nu_A(x) = 1$, for $\forall x \in X$, the intuitionistic fuzzy set A is reduced to an ordinary fuzzy set [9].

Definition 4. [14] Suppose there are two intuitionistic fuzzy numbers $\alpha = (\mu_\alpha, \nu_\alpha)$ and $\beta = (\mu_\beta, \nu_\beta)$, in which $\mu \in [0,1]$, $\nu \in [0,1]$, $\mu + \nu \leq 1$ and $\pi = 1 - \mu - \nu$. Let λ be a real number, then the arithmetic operations are defined as follows:

(1) Addition operation

$$\alpha + \beta = (\mu_\alpha + \mu_\beta - \mu_\alpha \mu_\beta, \nu_\alpha \nu_\beta); \tag{4}$$

(2) Multiplication by a real number operation

$$\lambda\alpha = (1 - (1 - \mu_\alpha)^\lambda, \nu_\alpha^\lambda), \lambda > 0. \tag{5}$$

Definition 5. [15] Let $\alpha = (\mu_\alpha, \nu_\alpha)$ and $\beta = (\mu_\beta, \nu_\beta)$ be two intuitionistic fuzzy numbers, then projection of the α on β is defined as follow:

$$Proj_\beta(\alpha) = \frac{\alpha \cdot \beta}{|\beta|} = \frac{\mu_\alpha \mu_\beta + \nu_\alpha \nu_\beta + \pi_\alpha \pi_\beta}{\sqrt{\mu_\beta^2 + \nu_\beta^2 + \pi_\beta^2}}. \tag{6}$$

In general, the larger the value of $Proj_\beta(\alpha)$, the more the degree of the α approaching the β.

Definition 6. [15] Let $A = (\alpha_{ij})_{mn}$ and $B = (\beta_{ij})_{mn}$ are two intuitionistic fuzzy matrices. Then the projection of A on B is defined as follows:

$$Proj_B(A) = \frac{\sum_{i=1}^m \sum_{j=1}^n (\mu_{ij}^\alpha \mu_{ij}^\beta + \nu_{ij}^\alpha \nu_{ij}^\beta + \pi_{ij}^\alpha \pi_{ij}^\beta)}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n \left((\mu_{ij}^\beta)^2 + (\nu_{ij}^\beta)^2 + (\pi_{ij}^\beta)^2\right)}}. \tag{7}$$

3 A Framework for MCGDM

The MCGDM problems have some common characteristics: multiple alternatives, multiple criteria and multiple decision makers. The aim of this kind of decision is to screen and rank the alternatives based on the preferences expressed by decision makers. Due to different education background and subjective perception, different decision makers have difference importance during decision-making process. Besides, it is quite common that there is no decision makers' weights information for the MCGDM, even decision makers' weights are decisive. Therefore, we construct a new framework to deal with MCGDM problems without any decision makers' weights information, as shown in Fig. 1.

Suppose that there are m alternatives $A_i, i \in \{1, 2, \cdots, m\}$ to be ranked. The performances of n criteria $C_j, j \in \{1, 2, \cdots, n\}$ are defined in linguistic terms which are obtained from t decision makers $D_k, k \in \{1, 2, \cdots, t\}$. Each decision

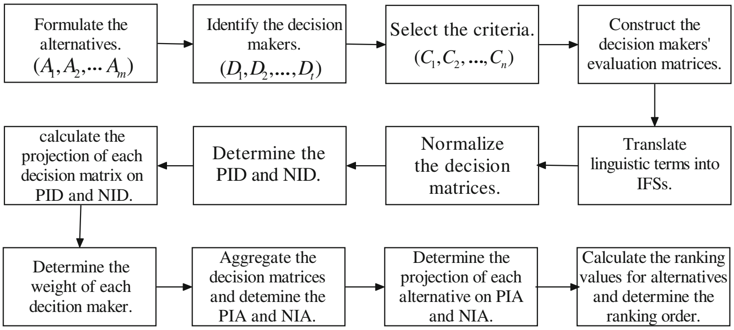


Fig. 1. The framework of proposed method

maker expresses his preference to each alternative regarding to each criterion using linguistic terms. Then decision matrices are constructed based on the all preferences, shown as follows:

$$D^k = \begin{pmatrix} r_{11}^k & r_{12}^k & \cdots & r_{1n}^k \\ r_{21}^k & r_{22}^k & \cdots & r_{2n}^k \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1}^k & r_{m2}^k & \cdots & r_{mn}^k \end{pmatrix}. \tag{8}$$

3.1 Decision Makers' Weights Determination

TOPSIS, which was first developed by Hwang and Yoon [5], is a technique for establishing the order of preference by its similarity to the ideal point [13]. Inspired by [13,16], we present an extended TOPSIS method for the decision makers' weight determination under an intuitionistic fuzzy environment.

First, we construct the positive ideal decision (PID) matrix D^+ based on the average of all decision matrices, which is treated as a cursory group consensus of decision-making group. The equation is as follows:

$$D^+ = \begin{pmatrix} r_{11}^+ & r_{12}^+ & \cdots & r_{1n}^+ \\ r_{21}^+ & r_{22}^+ & \cdots & r_{2n}^+ \\ \vdots & \vdots & & \vdots \\ r_{m1}^+ & r_{m2}^+ & \cdots & r_{mn}^+ \end{pmatrix}, \tag{9}$$

where

$$r_{ij}^+ = \frac{1}{t} \sum_{k=1}^{t} r_{ij}^k. \tag{10}$$

Then, the negative ideal decision (NID) matrix D^- is determined as the most indeterminate decision matrix, which means that the indeterminacy degree of worst decision matrix possesses is the highest in all decision matrices. The equation is as follows:

$$D^- = \begin{pmatrix} r_{11}^- & r_{12}^- & \cdots & r_{1n}^- \\ r_{21}^- & r_{22}^- & \cdots & r_{2n}^- \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1}^- & r_{m2}^- & \cdots & r_{mn}^- \end{pmatrix}, \tag{11}$$

where

$$r_{ij}^- = (\min_t \mu_{ij}^k, \max_t \nu_{ij}^k). \tag{12}$$

Since PID and NID have been determined, we consider the projection of each decision matrix on the ideal decision matrices. According to Eq. (7), the projection of D^k on D^+ can be calculated as follows:

$$\mathrm{Proj}_{D^+}(D^k) = \frac{\sum_{i=1}^m \sum_{j=1}^n (\mu_{ij}^{H^k} \mu_{ij}^{H^+} + \nu_{ij}^{H^k} \nu_{ij}^{H^+} + \pi_{ij}^{H^k} \pi_{ij}^{H^+})}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n \left((\mu_{ij}^{H^+})^2 + (\nu_{ij}^{H^+})^2 + (\pi_{ij}^{H^+})^2\right)}}. \tag{13}$$

The projection of D^k on D^- can be calculated as follows:

$$\mathrm{Proj}_{D^-}(D^k) = \frac{\sum_{i=1}^m \sum_{j=1}^n (\mu_{ij}^{H^k} \mu_{ij}^{H^-} + \nu_{ij}^{H^k} \nu_{ij}^{H^-} + \pi_{ij}^{H^k} \pi_{ij}^{H^-})}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n \left((\mu_{ij}^{H^-})^2 + (\nu_{ij}^{H^-})^2 + (\pi_{ij}^{H^-})^2\right)}}. \tag{14}$$

After that, the relative closeness of each individual decision matrix is determined as follows:

$$RC^k = \frac{\mathrm{Proj}_{D^-}(D^k)}{\mathrm{Proj}_{D^-}(D^k) + \mathrm{Proj}_{D^+}(D^k)}. \tag{15}$$

The relative closeness is used to determine the rank order for all decision makers. Obviously, the closer to the best decision result, the larger the value of RC^k is. Therefore, based on the relative closeness, the weight of each decision maker is obtained as follows:

$$\varpi_k = \frac{RC^k}{\sum_{k=1}^t RC^k} \tag{16}$$

where ϖ_k represents the wight of the kth decision maker, and $\varpi_k \geq 0$, $\sum_{k=1}^t \varpi_k = 1$. We use the vector $\varpi = (\varpi_1, \varpi_2, \cdots, \varpi_t)$ to represent the decision maker's weight vector.

3.2 Alternatives Ranking Order Determination

As the weights for decision makers are determined, the aggregated decision matrix can be calculated as follows:

$$D = (D^1, D^2, \cdots, D^t) \cdot (\varpi_1, \varpi_2, \cdots, \varpi_t)^T$$
$$= \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{pmatrix}, \tag{17}$$

where r_{ij} is the group preference IFS of alternative A_i respect to attribute C_j and

$$r_{ij} = \sum_{k=1}^{t} \varpi_k \cdot r_{ij}^k. \tag{18}$$

Then, the positive ideal alternative is determined as follows:

$$A^+ = (r_{11}^+, r_{12}^+, \cdots, r_{1n}^+), \tag{19}$$

where

$$r_{ij}^+ = (\max_m \mu_{ij}, \min_m \nu_{ij}). \tag{20}$$

The positive ideal alternative (PIA) should has the maximal separation from the PIA A^+. It is a very natural idea that we consider following decision:

$$A^- = (r_{11}^-, r_{12}^-, \cdots, r_{1n}^-), \tag{21}$$

where

$$r_{ij}^- = (\min_m \mu_{ij}, \max_m \nu_{ij}). \tag{22}$$

Then, the projection of each alternative on the ideal alternatives are calculated as follows:

$$\mathrm{Proj}_{A^+}(A^i) = \frac{\sum_{j=1}^n (\mu_{ij}^{A^i}\mu_{ij}^{A^+} + \nu_{ij}^{A^i}\nu_{ij}^{A^+} + \pi_{ij}^{A^i}\pi_{ij}^{A^+})}{\sqrt{\sum_{j=1}^n \left((\mu_{ij}^{A^+})^2 + (\nu_{ij}^{A^+})^2 + (\pi_{ij}^{A^+})^2\right)}}, \tag{23}$$

$$\mathrm{Proj}_{A^-}(A^i) = \frac{\sum_{j=1}^n (\mu_{ij}^{A^i}\mu_{ij}^{A^-} + \nu_{ij}^{A^i}\nu_{ij}^{A^-} + \pi_{ij}^{A^i}\pi_{ij}^{A^-})}{\sqrt{\sum_{j=1}^n \left((\mu_{ij}^{A^-})^2 + (\nu_{ij}^{A^-})^2 + (\pi_{ij}^{A^-})^2\right)}}. \tag{24}$$

Finally, the relative closeness is defined to determine the rank order for all alternatives. We have

$$RV_i = \frac{\mathrm{Proj}_{A^-}(A^i)}{\mathrm{Proj}_{A^+}(A^i) + \mathrm{Proj}_{A^-}(A^i)} \tag{25}$$

It is obvious that the bigger RV_i, the better alternative A^i is.

4 Site Selection for PFEVCS

PFEVCS site selection, as the preliminary work of EVCS construction, is quite important in the whole life cycle, which has significant impacts on the service quality and operational efficiency of EVCS. In this section, the proposed method is employed to select the optimal EVCS site from all the alternatives. In order to determine the possible alternatives and obtain the linguistic preference ratings, three groups of expert panels ($k = 1, 2, 3$) with expertise in the fields of environment, economy, society were formed. After reviewing the project feasibility research reports, the expert panels finally selected four EVCS sites as alternatives A_1, A_2, A_3 and A_4. EVCS sites alternatives A_1, A_2, A_3 and A_4 are respectively located at Wuhou district, Qingyang district, Jinniu district and Chenghua district in Chengdu.

The MCDM problem related to EVCS site selection includes three criteria (environment, economy and society) and eight sub-criteria ($j = 1, 2, \cdots, 8$). After reviewing the general information of all the alternatives, each expert group gave the linguistic ratings judgments for the sub-criteria performance of each alternative, and the results are listed in Table 1. According to the transform rules in Table 2, we transform the linguistic terms into intuitionistic fuzzy sets.

Step 1. We transform the decision matrices into intuitionistic fuzzy sets, and then determine the PID and NID respectively based on Eqs. (9)–(12). The projection of each decision matrix on PID and NID are calculated using Eqs. (13), (14). Then, according to Eq. (15), (16), the weights for decision makers are calculated as $\varpi = (0.3351, 0.3317, 0.3332)$.

Table 1. The decision matrix of decision makers

Decision makers	Alternatives	Criteria							
		C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
D_1	A_1	MH	ML	H	VH	H	VH	ML	L
	A_2	ML	VL	M	M	MH	M	VL	M
	A_3	M	L	M	L	M	H	L	MH
	A_4	ML	MH	VL	MH	L	VL	MH	VL
D_2	A_1	H	L	VH	H	H	H	L	L
	A_2	M	M	ML	L	VH	H	L	ML
	A_3	MH	L	M	H	M	MH	ML	MH
	A_4	L	VH	L	H	ML	L	H	VL
D_3	A_1	H	L	H	MH	VH	H	ML	L
	A_2	ML	ML	MH	ML	MH	VH	M	H
	A_3	M	L	L	L	M	M	M	MH
	A_4	L	VH	ML	H	MH	VL	L	VL

Table 2. The relationship between IFSs and linguistic terms [12]

Linguistic terms	Corresponding IFNs
Very low (VL)	[0.15,0.80]
Low (L)	[0.25,0.65]
Medium low (ML)	[0.35,0.55]
Medium (M)	[0.5.0.4]
Medium high(MH)	[0.65,0.25]
High (H)	[0.75,0.15]
Very high (VH)	[0.85,0.10]

Step 2. Based on the weights for decision makers and Eq. (17), the aggregated decision matrix is calculated. Besides, the PIA and NIA are determined, respectively, using Eqs. (19)–(22).

Step 3. The projection of each alternative on PIA and NIA are calculated based on Eqs. (23), (24). Then the ranking value for each alternative is calculated as $RV = (0.9583, 0.9445, 0.9344, 0.8973)$. Therefore, the ranking order for alternatives is $A_1 > A_2 > A_3 > A_4$.

As a result, site A_1 is the best site for public fast electric vehicle charging station.

5 Conclusion

In MAGDM problems, it is quite common that decision makers' weight information is partly or completely unknown because of such problems as decision makers lack of knowledge, the complexity of the decision-making environment, and scientific development limitations. Therefore, it is important to research a solution to MAGDM problems when no weight information is available. In this paper, we applied the intuitionistic fuzzy projection-based approach to determine the optimal site for public fast electric vehicle charging station to promote the popularization of EVs in Chengdu. In our model, the extended TOPSIS method under an intuitionistic fuzzy environment was developed to determine the attributes weights. With the weight determination method proposed in this paper, we effectively reduced the subjective effects to a minimum, so, because less information was required to make a decision using the proposed method, we demonstrated how MCGDM problems may become easier based on our approach. Finally, under an intuitionistic fuzzy environment, the proposed approach was shown to be able to handle MCGDM problems effectively, thus guaranteeing that there is less deviation in the decision making.

Acknowledgements. This work was supported by the National Natural Science Foundation of China under the Grant No. 71671118.

References

1. Atanassov KT (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20(1):87–96
2. Dyer JS, Fishburn PC et al (1992) Multiple criteria decision making, multiattribute utility theory: the next ten years. Manage Sci 38(5):645–654
3. Frade I, Ribeiro A et al (2011) Optimal location of charging stations for electric vehicles in a neighborhood in Lisbon, Portugal. Transp Res Rec J Transp Res Board 2252(2252):91–98
4. Guo S, Zhao H, Yan J (2015) Optimal site selection of electric vehicle charging station by using fuzzy topsis based on sustainability perspective. Appl Energy 158:390–402
5. Hwang C, Yoon K (1981) Multiple attribute decision making. Springer, Heidelberg
6. Li S, Huang Y, Mason SJ (2016) A multi-period optimization model for the deployment of public electric vehicle charging stations on network. Transp Res Part C Emerg Technol 65:128–143
7. Shareef H, Islam MM, Mohamed A (2016) A review of the stage-of-the-art charging technologies, placement methodologies, and impacts of electric vehicles. Renew Sustain Energy Rev 64:403–420
8. Sheppard CJR, Gopal AR et al (2016) Cost-effective electric vehicle charging infrastructure siting for Delhi. Environ Res Lett 11(6):064,010
9. Shu MH, Cheng CH, Chang JR (2006) Using intuitionistic fuzzy sets for fault-tree analysis on printed circuit board assembly. Microelectron Reliab 46(12):2139–2148
10. Wang JQ, Wang J et al (2014) An outranking approach for multi-criteria decision-making with hesitant fuzzy linguistic term sets. Inf Sci 280:338–351
11. Wang JQ, Wang P et al (2015) Atanassov's interval-valued intuitionistic linguistic multicriteria group decision-making method based on the trapezium cloud model. IEEE Trans Fuzzy Syst 23(3):542–554
12. Wang T, Liu J et al (2016) An integrating owactopsis framework in intuitionistic fuzzy settings for multiple attribute decision making. Comput Ind Eng 98:185–194
13. Wu ZB, Zhong L (2016) Weight determination for magdm with linguistic information based on it2 fuzzy sets. In: IEEE International Conference on Fuzzy Systems, pp 880–887
14. Xu ZS (2007) Intuitionistic fuzzy aggregation operators. IEEE Trans Fuzzy Syst 15(6):1179–1187
15. Yue Z (2013) An intuitionistic fuzzy projection-based approach for partner selection. Appl Math Model 37(23):9538–9551
16. Yue ZL (2011) An extended topsis for determining weights of decision makers with interval numbers. Knowl Based Syst 24(1):146–153
17. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353
18. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning. Inf Sci 8(4):301–357

If the Medical Expenses Do Effect the Rural Residents' Consume

Songtao Jiang[1(✉)] and Yuan Ji[2]

[1] School of Economics, Sichuan University,
Chengdu 610064, People's Republic of China
jst714@qq.com

[2] Sichuan Technology and Business University,
Chengdu 611745, People's Republic of China

Abstract. We exploit a fixed effect model to examine if the uncertain medical expenses of rural residents reduce their total consume using the provincial-level panel data from 2000–2013. To obtain a robust results, we then add the model with the time fixed effect. The results show that, rural residents' medical expenses share raise will reduce the share of rural residents' total consume in GDP by more than 6%, and this effect is more obvious than the rural residents' education expenses which is less than 6%. We also find that if the government increase the funds proportion of health care, the medical demand of rural residents will raise which may reduce the percentage of rural residents' total consume in GDP by more than 6%. By the way, liquidity constrains may also enhanced the rural residents' precautionary saving motives based on the medical expenses. So we can conclude that the medical expenses of rural residents do effect the consume demand of rural households significantly. And the government need to control the rural residents' medical expenses if they attempt to expand the consumption of rural residents.

Keywords: Medical expenses · Rural residents consume · Panel data · Saving motives

1 Introduction

As the uncertainty of global economic recovery increases, the major economic growth engines of investment and export are facing the constraints of environmental protection and trade barriers. In order to keep economic growth at a reasonable rate China must pay more attention to the role of consumption in stimulating the economy.

Figure 1 illustrates the change in consumption and investment rates between 1978 and 2013 in the gross domestic product of the expenditure approach. Where S represents the consumption rate, and I represents the investment rate. It is clearly that from 1978 to 2013 the rate of consumption in our country shows a general downward trend, especially after 2000, it shows a steep downward

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_143

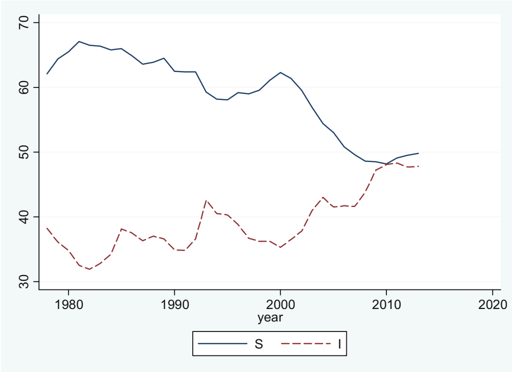


Fig. 1. Trends of consumption and investment rates in 1978–2013

trend, and it does not start to stabilize until 2010, and rises slightly. In sharp contrast to this is that since 1978, China's economic growth into the fast-rising channel, which indicated that China's consumption rate is not with China's rapid economic growth and increase.

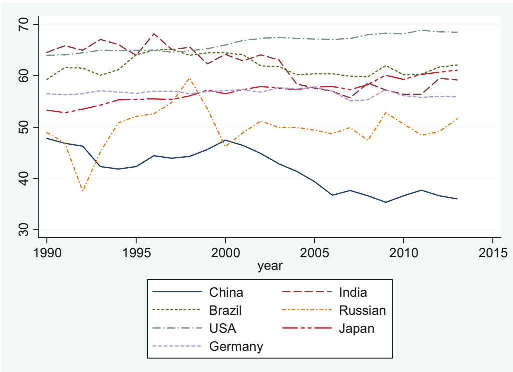


Fig. 2. International comparison of final consumption proportion of residents in GDP

From the international comparison point of view, we can also get the same conclusion. Figure 2 shows the trend of final consumption among main countries in GDP. Can be seen from Fig. 2, China's consumption rate is not only lower than developed countries such as European countries and the United States, with India, Brazil and other emerging countries, which were also significantly lower. Developed countries from Europe and the United States, its consumption rate is generally high, and long-term stable operation. From the Fig. 2, The US consumption rate has maintained a high level of smooth operation (basically remained above 64), and in a steady rise with a slight trend; Germany and Japan, similar to the situation, although the final consumption rate ate not as high as the United States, but can also maintained at a high level (basically remained around 55) and with a more stable trend. While the emerging market countries, the final consumption rate of residents compared with the developed countries is characterized by the fluctuations are generally large, including China. Whether China, India, Brazil or Russia, the consumption rate of its residents in the 1990–2013 years have experienced a more vibration. It is obvious that the final consumption rate of Chinese residents is still very low compared with that of emerging market countries.

When looking at the household consumption which play a decisive role of consumption, we can find that the rural residents' consumption as a share of household consumption has not increased with the urban residents', and it makes rural consumption become a short board of China's strategy of expanding domestic demand.

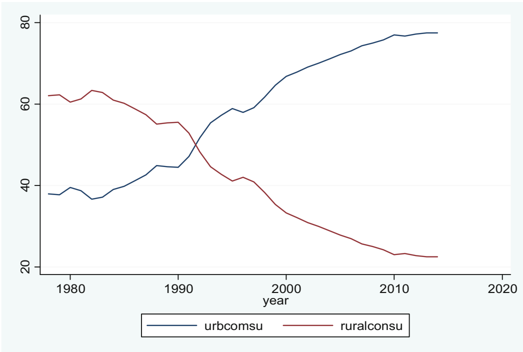


Fig. 3. The share of urban and rural household consumption in 1978–2013

Figure 3 shows the proportion of urban residents' consumption and rural residents' consumption in China's household consumption during 1978–1930, where urbcomsu' represents urban residents consumption in the proportion of consumption, 'ruralconsu' represents rural residents consumption in the proportion of consumption. As can be seen from the figure, before 1992, the proportion of consumption of rural residents in China's household consumption in a dominant position, but the early 90s of last century, the proportion of rural residents began to decline sharply and the urban residents began to dominate the proportion of consumer demand. The consumption gap between urban and rural residents has been widening year by year, until 2010, the gap widening trend has slowed down.

Generally, there are many factor which constrain China' rural consumption, but the imperfect medical security system and the uncertainty of medical expenditure of rural resident cannot be ignored. As the medical consumption has the characteristics of uncertainty and urgency, the impact of medical consumption on rural residents' overall consumption may be significant. This paper using the provincial-level panel data from 2000–2013 to test if the uncertain medical expenses of rural residents reduce their total consume by adopting fixed effect model. Through the empirical research, we expecting find out the relationship between medical expenses of rural residents and rural residents' total consumption. And based on the conclusion, we may provide some reference for the policy of rural residents' consumption.

2 Literature Review

Scholars have studied the reasons for the downturn in rural consumption and how to start rural consumption from different aspects: Wang's research [14] indicated that the uncertainty expectations of rural residents are one important reason that restricts the consumption in rural areas in China. Similar views also include Tan [11]. Qu [13] found that the pessimistic expectations on future expenditure of rural residents will place greater emphasis on deposit but not expenditure. Therefore to start rural consumption China's policy makers should improve the rural consumption environment and reduce the burden on farmers. Sun and Zhu' research [9] also emphasized the importance of optimizing the rural consumption environment to improve the consumption of rural residents. Cai [1] had point that in addition to income levels, rural social security system construction is also an important factor affection rural consumption. Clearly, as a important part of social security system, medical security system are closely related to medical expenditure, which indicate that the medical expenditure of rural residents may play an important role in consumer decision making.

There are also scholars' studies focused on the consumer environment and the expectations of rural residents. Liu [7] found that insufficient financial support limited the development of rural residents' consumption. Yang and Chen's research [16] based on the view of rural public goods supply, proposed that increasing rural public goods supply can reduce the uncertainty of expenditure expectations of rural residents which may benefit the rural consumption. Zhou and Yang analyzed the consumption both urban and rural areas and pointed out that, precautionary saving motive may also effect rural residents as well as urban residents.

Furthermore, some scholars studied about rural medical expenditure and the existing literature mainly focuses on two aspects: first, the impact factors of rural residents' medical expenditure such as Tan and Zhang et al. [10] whose study explores the influencing factors of rural households' medical consumption spending and its demand elasticity based on QUAIDS model and the cross-section; And Huang and Liu [6] provided a similar study, which state that the income level, government health investment and other factors do affect the medical consumption of rural residents; Zhao [17] used the 1993–2012 panel data to analysis the regional differences of rural residents' health care consumption, and state that the income, education, economic development and other factors on the impact of rural resident's medical spending are significant regional differences. Second, the function of rural medical security system such as Bai and Li et al. [3] found that the insurance coverage increases non-health care consumption by more than 5% and the insurance effect exists even for households with medical spending, by exploiting a quasi-natural experiment; Ding and Ying et al. [2]

also stated that the propensity to consume of rural households is universally low in China, and the medical insurance do increased rural durables consumption significantly, base on the analysis of China Health and Nutrition Survey panel data from 1991 to 2009. Tan and Qin [12] used IV-TOBIT model to test the crowding out effect of China's households' medical expenses and conclude that the medical expenses do reduce the households' expenses on food and clothing, whose research also used the China Health and Nutrition Survey data (2013). Although the above study about rural medical expenditure focused on difference aspects, they provided us with the method of reference.

In general, although many studies focused on the rural residents' consumption expansion directly or indirectly mentioned factors such as rural residents' expectation, consumption environment, and precautionary saving motive. But there are few research put rural residents' medical expenses as an independent factor which effect the rural residents' consumption behavior. For this reason, we develop an empirical model that includes both residents' medical expenses and a set of factors that may affect the rural residents' medical expenses. This model also contain standard economic variables. Our study adds to the current literature by utilizing pooled time series and provincial cross-sectional data over a longer period than the household survey data cover. In addition, by utilizing medical expenses variable, we are looking for if the medical expenses of rural residents do effect the rural consume.

3 Investigation of if Medical Expenses Affect Rural Residents' Aggregate Consumption

The equation to be estimated for (i) provinces and (t) time periods is as follows:

$$\left(\frac{C}{Y}\right)_{it} = \beta_0 + \beta_1 \mathrm{medratio}_{it_{it}} + \beta_2 \mathrm{goeratio}_{it} + \beta_3 \mathrm{loandep}_{it} + \gamma_1 P_1 + \cdots + \gamma_{i-1} P_{i-1} + \lambda_1 D_1 + \cdots + \lambda_{t-1} D_{t-1} + \sum \sigma_i Z_{it} + \mu_{it},$$

where

$\left(\frac{C}{Y}\right)_{it}$: rural residents' consumption as ashare of provincial domestic product;

$\mathrm{medratio}_{it}$: the share of rural household medical expense in rural household consumption per capita;

$\mathrm{goeratio}_{it}$: local government expenditure on health as a share of total local government expenditure;

$\mathrm{loandep}_{it}$: the share of loans plus deposits in provincial domestic product;

P_i, D_t : provincial and time dummies;

Z_{it} : vectors of economic system variables;

growth_{it} : the annual growth in provincial domestic product;

dinco_{it} : tprovincial net income of rural residents per capita;

$\mathrm{depratio}_{it}$: dependency ratios;

$\mathrm{eduratio}_{it}$: the share of rural household education expense in rural household consumption per capita.

This empirical model used here combines standard consumption determinants relating to income, government expenditure and family structure which expected to influence households' consumption decisions. While this model also concentrated to rural residents' medical expenses and the medical expenses expectation. We use the share of rural residents' consumption in domestic product $\left(\frac{C}{Y}\right)_{it}$ as our dependent variable. This is the share that appears to behave differently compared to urban. Our panel data set includes provincial data for 31 provinces between 2000 and 2013. The variation across provinces allows us to explore possible factors in addition to the standard consumption variables.

The impact of medical expenses on household's consumption is manifest in two aspects: First, medical spending will be squeezed the share of other consumption under the total budget constraints. Second, medical expenditure will increase people's cautious motivations, which will increase precautionary savings. In addition, as the medical treatment has high uncertainty, residents could not predict the medical expenses in the time and amount, which may strengthen residents' saving motive. To test if this impact do affected China's rural consumption we applied the share of rural household medical expense in rural household consumption per capita as a key variable ($\mathrm{medratio}_{it}$). A reasonable hypothesis is that the increase in the share of rural household medical expenses in rural household consumption per capita will lead rural residents to reduce consumption.

A second factor that might influence rural residents' consumption relating to medical expenses is government expenditure on health ($\mathrm{goeratio}_{it}$). If the government is taking care of residents' health care needs, households would not need to save for the possible expenditures or may save less. So we can hypothesis that the government health expenditure variable is positively correlated with the dependent variable.

The third factor we concerned relating to the household medical expenses and the rural residents' consumption is the ability of rural residents to use the financial system to smooth consumption over time. So we adopt financial development indicators as the third variable which can lessened liquidity constraints of rural residents. The standard financial development indicators are the percentage of total financial assets to GDP [4] or money supply measures such as M2 as a share of GDP [8]. But we do not have access to financial assets data for all provinces for the period of time we study, and the money supply measures do not vary by province but only vary by country. Therefore we follow [18] and use the total deposits and loans issued by all financial institutions as a percent of GDP reported by each province ($\mathrm{loandep}_{it}$) as a proxy variable for liquidity constraint.

For the control variables in the Z vector, we include a number of factors raised by the literature: (growth_{it}) is the annual growth in provincial domestic product to interpret the economic development of the provinces; (dinco_{it}) is provincial net income of rural residents per capita; ($\mathrm{depratio}_{it}$) is the provincial dependency ratios, in principle, higher dependency ratio will lead to higher consumer spending. In the empirical model we also adopt the share of rural household education expense in rural household consumption per capita as a control variable

($eduratio_{it}$) because it has been pointed out by the literature that education expenses is an important factor that affected residents saving motions. By the way, adopting the education expenses variable can make a comparison with the medical expenses variable.

4 Estimation and Results

Table 1 reports our descriptive statistics as well as the between-province standard deviation and within-province standard deviation (over the entire time period) of our main variables in our model. This table confirms the considerable variations of the dependent and independent variables of interest in our data set.

We begin our estimation with a pooled OLS model. The result reported in Table 2 model 1. This model is estimated with pooled ordinary least squares using

Table 1. Descriptive statistics

Variable		Mean	Std. Dev	Min	Max	Observations
$\left(\frac{C}{Y}\right)_{it}$	Overall	0.1237	0.0665	0.0185	0.3669	$N = 434$
	Between		0.0485	0.0285	0.2252	$n = 31$
	Within		0.0463	0.034	0.2838	$T = 14$
medratio	Overall	0.07	0.0219	0.0144	0.1399	$N = 434$
	Between		0.0168	0.0252	0.1008	$n = 31$
	Within		0.0144	0.0353	0.1376	$T = 14$
loandep	Overall	2.5031	0.8581	1.28	6.66	$N = 434$
	Between		0.8176	1.5621	5.9279	$n = 31$
	Within		0.2966	1.3603	4.1188	$T = 14$
eduratio	Overall	0.095	0.0297	0.0099	0.1841	$N = 434$
	Between		0.0226	0.0216	0.1314	$n = 31$
	Within		0.0197	0.0473	0.1541	$T = 14$
depratio	Overall	38.0565	7.3399	19.27	57.58	$N = 434$
	Between		6.3396	25.2482	51.7112	$n = 31$
	Within		3.8588	28.8267	52.9967	$T = 14$
goeratio	Overall	0.0215	0.0091	0.0019	0.0488	$N = 434$
	Between		0.0078	0.0032	0.0403	$n = 31$
	Within		0.0049	0.0071	0.0383	$T = 14$
dinco	Overall	4073.121	2348.768	1332.142	14691.7	$N = 434$
	Between		1790.959	2137.623	9200.569	$n = 31$
	Within		1550.96	131.1782	9564.254	$T = 14$

The most data used in this paper are from China Statistical Yearbook. The provincial net income of rural residents per capita data were adjusted based on 1999 CPI. The government expenditure on health and education data are budgetary data from both China Statistical Yearbook and China Fiscal Yearbook.

Table 2. Estimation results

	Model 1	Model 2	Model 3	Model 4	Model 15
edratio	−0.635*** (0.094)	−0.643*** (0.124)	−0.643*** (0.169)	−0.390*** (0.098)	−0.223* (0.117)
goeratio	1.171*** (0.261)	−0.674** (0.290)	−0.674* (0.385)	−0.786*** (0.187)	-
loandep	−0.007*** (0.002)	−0.017*** (0.005)	−0.017* (0.010)	−0.026*** (0.008)	-
eduratio	0.479*** (0.067)	0.586*** (0.103)	0.586*** (0.136)	0.389*** (0.068)	0.321*** (0.058)
beds	- -	- -	- -	- -	−0.000*** (0.000)
save	- -	- -	- -	- -	−0.071*** (0.014)
yr00_04	-	- -	- -	0.055*** (0.011)	0.044*** (0.009)
yr05_09	- -	- -	- -	0.004 (0.006)	−0.006 (0.004)
growth	−0.008*** (0.001)	−0.009*** (0.001)	−0.009*** (0.001)	−0.005*** (0.001)	−0.005*** (0.001)
depratio	0.003*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.001* (0.000)	0.001*** (0.000)
dinco	−0.000*** (0.000)	−0.000*** (0.000)	−0.000*** (0.000)	0.000* (0.000)	−0.000*** (0.000)
_cons	0.905*** (0.090)	1.152*** (0.083)	1.152*** (0.181)	0.745*** (0.158)	0.698*** (0.147)
F test all u_i = 0	$Prob > F = 0.00$				
N	434	434	434	434	406
adj. R^2	0.741	0.689	-	-	-
R^2	0.745	0.716	-	-	-
Within-R^2	-	0.716	0.716	0.808	0.823

Standard errors in parentheses
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

the full data set from 2000–2013. The share of rural household medical expense variable (medratio$_{it}$) is significant and negative. Which met our hypothesis that the increase in the share of rural household medical expenses in rural household consumption will lead rural residents to reduce consumption. While the government expenditure on health variable (rugeoratio$_{it}$) is significant and positive which met our hypothesis too that the government health expenditure variable is positively correlated with the dependent variable.

There after we add cross-provincial fixed effect to controls for omitted variables that vary across provinces but do not change over time. Furthermore, vast differences in economic, industrial and fiscal structures across provinces are likely to influence household consumption. So the intercepts generated for each province in this model should absorb the influences of these omitted

Table 3. Robustness checks

Tests	Methods	Hypothesis	Results
Heteroskedasticity	Modified Wald test	H0: sigma$(i)^2$ = sigma^2 for all i	Prob > chi2 =0.0000
Cross-section correlation	Frees' test	H0: cross-sectional independence	5.828 > alpha = 0.01 : 0.36
Sequence correlation	Wooldridge test for autocorrelation in panel data	H0: no first order autocorrelation	$Prob > F = 0.0000$

variables that differ for each province but are constant over time. Model 2 is also estimated with ordinary least squares, but with provincial fixed effects reported in Table 2. The F-statistics of the test on the significance of provincial fixed effects show that there are significant variations among provinces ($F(30, 396) = 11.18, \Pr ob > F = 0.0$). In model 2 the government expenditure on health variable (rugeoratio$_{it}$) is significant but negative, reflecting that government expenditure on health increasing would not stimulate the rural residents' consumption. Although the share of rural household medical expense variable (medratio$_{it}$) is also significant and negative.

Table 3 presents a series of robustness checks on model 2 including heteroskedasticity, cross-section correlation, and sequence correlation. As the test results shown in Table 3. The result estimated in model 2 is not robustness. So we re-estimates model 2 using Daniel Hoechle's [5] method to get stronger results in model 3. Compared with model 2, the estimate result in model 3 are fairly similar, only the significant level of the coefficients has changed. But our main independent variables are still significant.

Based on model 3, we then add time-fixed effects to control for variables that are constant across provinces but evolve over time, such as the introduction of new healthcare regime "Xinnong he" (New Cooperative Medical Scheme), the nationwide economic reforms deepen in China rural area, etc. Model 4 replicates model 3 except that these estimates include the time fixed effects as well as the provincial fixed effects. We included dummy variables chosen for several time periods to capture the trend of the introduction of new healthcare regime. As the New Cooperative Medical Scheme start the pilot at 2003 and realized full coverage at 2009, considering the lag of policy, our first time dummy starts from 2000 and ends in 2004. The next time dummy covers the period of 2005 to 2009. The last dummy variable covers the time period between 2010 and 2013. In model 4, at least one newly generated time dummy variables is statistically significant and produce coefficients with a trend: over time the coefficient increase each year. The basic findings of model 3 remain robust in model 4, although some coefficient changed in magnitude. So we can draw our conclusion based on model 4.

In model 4, our main dependency variable were all significant although we added the time-fixed effects. The share of rural household medical expense is negative, which is consistent with our assumption. The government expenditure on health is negative, implying that government expenditure on health do not release rural residents' saving motive, which doesn't met our hypothesis. The financial development is negative in model 1–model 4, which indicated that the financial development did not lessened rural households' liquidity constraint. In Table 2 we also found that the shares of rural household education expense in rural household consumption, which we adopted to make a comparison with the medical expenses as a control variable, were significant and positive in model 1–model 4. So we can conclude that the impacts of medical expenditure and education expenditure on rural residents' consumption are not consistent.

As the government expenditure on health and the financial development's estimated results did not consistent with our expectations and the general theory, we re-estimated model 4 by alternative measures of these two variables as a further robustness check, presented in model 5.

In model 5 we used the number of beds in township health institutions to replace the government expenditure on health. Simultaneously, we adopted the net saving rate of rural residents [15] per capita to replace the financial developments as the proxy variable of liquidity constraint of rural residents. Since the statistical agency is no longer announced the number of beds in township health institutions of Shanghai and Beijing after 2009, our estimation in model 5 did not include the two sections of Shanghai and Beijing. Also because we adopted the panel data contain 31 sections, so there is no substantial impact on the estimation from the elimination of the two sections.

In model 5, although there are some changes in the estimation results compared with model 4, but the share of rural household medical expense remain significant and negative, by the way the shares of rural household education expense are also significant and positive, which are corresponded with model 4. Furthermore, the number of beds in township health institutions is significant and negative too, as a part of government expenditure on health, which indicated that the result of government expenditure on health estimated in model 4 has quite robustness. As the net saving rate of rural residents is significant and negative, we can conclude that the rural residents faced obvious liquidity constraints. In addition, as the model 5 adopt the net saving rate of rural residents, the coefficient of the share of rural household medical expense becomes smaller, indicated that medical expense always accompanied with liquidity constraint, which is associated with the amount of medical expenditure relatively large.

5 Discussion and Conclusion

We use various models and estimation techniques to examine if and how the medical expenses affect the variation in rural residents' consumption as a share of GDP. As our expectation the medical expenses do lessened the rural residents' consumption as a share of GDP. In our study, the most significant and robust discovery is that reducing rural residents' motivations for precautionary savings through government's expenditure at the local level have the opposite effect in increasing rural residents' consumption. While this result contrary to our expectation and the general theory, the alternative measure that captures the role of the government expenditure through the number of beds in township health institutions was also significant. From the direct meaning of this estimation result, we can conclude that the increasing of government spending on health has stimulated the rural household saving motive. But this is by no means a justification for reducing the fiscal responsibility of the government. On the contrary, the result implied, to some extent, that the government expenditure on rural health is not enough, and there is a heavy responsibility on the government to improve the rural medical and health environment. Because at the pooled medical and health environment level, the rural household's medical consumption demand has been repressed.

The financial development and its alternative measurer always robust regressors in all the models, while the measurements are small and negative, which means the liquidity constraints is a factor that do not bode well for the rural consumption. The shares of rural household education expense are always significant and positive, while the coefficient absolute value is large. As demonstrated in this study, this means that, the impacts of medical expenditure and education expenditure on rural residents' consumption are not consistent, so we need different options to develop the medical and education in rural areas.

For the control variables in the Z vector, although the annual growth in provincial domestic product is significant influenced the independent variable but the coefficient absolute value is too small, while the provincial net income of rural residents and the provincial dependency ratios suffer from the same problem.

These results point to several suggestions that the increasing medical expenses is the obstacle of boosting domestic consumption especially in rural areas. So the policy makers need to formulate policies to optimize the consumer environment in rural areas, just like increase investment in rural health care, although the government expenditure on health is negative in this paper. The cautious motive which lead rural household to increase their savings has been proved indirectly by the estimated result in this paper as the share of rural household medical expense in rural household consumption is significant and negative. In addition, as the liquidity constraint has also been found in the rural are as from the estimated results. To enhance the rural consumption, the government needs to not only lift the burden of rural households medical expenditure but lessen the liquidity constraint also. In our opinion, to increasing the coverage and funding levels of NRCMS is an effective method. Furthermore, increasing the accessibility and convenience of rural health care facilities may also a effective way.

References

1. Cai YZ (2009) Economic stimulus package and start-up of rural consumption in China: empirical study based on the decomposition of rural household income. J Financ Econ 9:4–12
2. Ding J, Ying M, Du Z (2013) China's rural household consumption behavior research: based on the perspective of health risk and medical insurance. J Financ Res 10:154–166

3. Ding JH, Ying ML, Du ZC (2013) Health insurance and consumption: evidence from China's new cooperative medical scheme. Consum Econ 10:154–166
4. Goldsmith RW (1969) Financial structure and development. Stud Comp Econ 70(4):31–45
5. Hoechle D (2007) Robust standard errors for panel regressions with cross-sectional dependence. Stata J 7(3):281–312
6. Huang XP, Liu H (2011) An analysis on the influencing factors of medical consumption of Chinese rural residents. Consum Econ 11:77–80
7. Liu GM (2011) Discussion and analysis about financial support to rural consumer market's exploitation. J Central Univ Financ Econ 6:35–40
8. Mckinnon RI (1973) Money and capital in economic development. Brookings institution
9. Sun HQ, Zhu C (2012) An empirical analysis of urbanization and rural consumption growth in China. Stat Decis 5:90–93
10. Tan T, Zhang YY, Jun HE (2014) An analysis of influencing factors and elasticity of rural households' medical consumption spending in China. J Shanghai Univ Financ Econ 3:63–69
11. Tan YH (2009) On fiscal policies of stimulating consumer demand in rural areas. J Central Univ Financ Econ 8:5–9
12. Tang Q, Qin XZ (2016) An empirical study on the crowding-out effect of family medical expenditure in China. Econ Sci 3:61–75
13. Tao Q (2009) Analysis on the consumption behavior and its constraint factors in rural China. Economist 9:54–60
14. Wang B (2012) The deep reason and improving ways of rural consumption demand in China. Consum Econ 1:29–32
15. Wang P (2014) Rural institutional change, liquidity constraint and consumption growth of rural residents in China. J Shanxi Univ Financ Econ 10:1–10
16. Yang L, Chen C (2013) An analysis of the effect of education and health public goods on rural residents' consumption: from the perspective of human capital promotion. J Agrotechnical Econ 9:4–12
17. Zhao ZR (2014) A study on reginal differences of influencing factors of rural residents' health consumption in China. Consum Econ 3:24–29
18. Zhou L, Wang Z (2008) An empirical analysis of financial development and economic growth in China. J Financ 10:1–13

A Discrete Time-Cost-Environment Trade-Off Problem with Multiple Projects: The Jinping-I Hydroelectric Station Project

Huan Zheng[1,2](✉)

[1] Management School, Chongqing Technology and Business University, Chongqing 400067, People's Republic of China
290871340@qq.com
[2] Engineering School, Widener University, Philadelphia 19013, USA

Abstract. Stronger environmental protection regulations and innovative contracting methods focused on reducing environmental impact through new incentives have placed increasing pressure on decision makers, especially those on large-scale construction projects. This increase has led to a need to identify optimal decision and select the most efficient construction modes for each activity to minimize not only the project's time and penalty costs but also its environmental impact. To address this need, this paper applies a mathematical model based on a hybrid genetic algorithm with fuzzy logic controller (flc-hGA) to the time-cost-environmental trade-off problem with multiple projects under a fuzzy environment. Applying the proposed method to the case study of the Jinping-I Hydropower Station, a large-scale project in Southwest China, clearly demonstrates its economic, technological, and social ecological effectiveness.

Keywords: Time-cost-environment · Trade-off · Multiple project · Fuzzyt · Hybrid genetic algorithm

1 Introduction

Two of the most crucial aspects of any such construction project are time and project cost, both of which have received considerable research attention [9,12]. One particularly important element of effective project scheduling theory and applications is the discrete time-cost trade-off problem (DTCTP) introduced by Harvey and Patterson [6]. Of more recent concern, however, are accusations that the construction industry is causing environmental problems that range from excessive consumption of global resources-both in terms of construction and operation-to pollution of the surrounding environment. Hydroelectric projects, particularly, contribute significantly to changes in river environments

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_144

[2,17]. Yet previous studies have paid more attention to evaluating the environmental impact at later stages of reservoirs and hydraulic electricity rather than considering it from the construction stage. Recently, however, a tightening of environmental protection regulations, especially for hydroelectric projects, has increased pressure on project managers to reduce the environmental impact through the selection of optimal construction modes. There is an imminent need, therefore, to study management decisions on selection methods to ensure more environmentally friendly means of construction during the project planning stage when the environmental impact can best be incorporated into other project objectives. For hydroelectric projects such as the JHS-I, this planning stage should include the optimization of construction and subsequent operations for environmental impact as well as time and cost.

The expanding scale of such construction projects worldwide, however, makes effective project management extremely complex. To deal with this complexity while still achieving management objectives, construction managers must employ a project management decision system that effectively controls total project duration, penalty costs, and environmental impact. One effective method for such control is the discrete time-cost-environment trade-off problem with multiple projects (DTCETP-mP), which is an extension of the original DTCTP. This analysis thus applies DTCETP-mP to solve the three main objectives of the JHS-I problem: (1) minimization of the total project duration, (2) minimization of the total project penalty cost, and (3) minimization of the environmental impact.

In non-routine projects such as the new construction at JHS-I [10], the duration of each activity and completion time may be uncertain, so the project manager must handle multiple conflicting goals in an uncertain environment in which information may be incomplete or unavailable. In this context, activity duration uncertainty can be modeled using either probability-based methods [12] or fuzzy set-based methods [18,19] depending on the situation and the project manager's preference. When a project manager has difficulty characterizing the random variables, as in the current scenario of a new construction project with unique activities and a lack of historical data, the fuzzy method is the most effective approach.

Because the DTCETP-mP is an extension of the DTCTP, it is an NP-hard problem that is difficult to solve [15]. Yet even the most exact of currently available methods can only solve small projects with under 60 activities, a far cry from the numerous activities and modes per activity involved in the large-scale, complex JHS-I project, whose optimal solutions are beyond the capabilities of traditional production scheduling methods like PERT (Program Evaluation and Review Technique) and CPM (Critical Path Method) [3]. More suitable heuristic solution procedures for solving the DTCETP-mP are thus needed, several of which have been suggested in the literature. For example, Franck et al. [4] demonstrated that a genetic algorithm (GA) performed slightly better than a tabu search (TS) procedure but required more computing. In earlier work, Wang et al. [16] addressed larger, more complex problems by introducing an improved hybrid genetic algorithm (hGA) that uses a fuzzy logic controller (flc) to adaptively regulate the GA parameters (including generation number, population size, crossover ratio, mutation ratio) and automate mutation ratio fine tuning. The effectiveness and efficiency of the proposed flc-hGA is evaluated through comparison with other methods.

This present application of the DTCETP-mP to identify optimal solutions for the JHS-I large-scale deeply buried tunnel group project amends the GA first introduced by Holland [7] to an flc-hGA. The analytic results indicate that this optimization method is both practical and highly efficient in solving the JHS-II problem and achieving minimization of the project's total duration, penalty costs, and environmental impact.

2 Problem Description and Mathematical Model

2.1 Problem Description

Because the JHS-I project is unique, lacking in historical data, technically difficult to construct, structurally complex, and subject to strict requirements and many construction uncertainties, it is difficult to estimate the exact duration of each activity or assign each a strict function [8,11]. Duration is thus characterized by fuzziness and subjective uncertainty [18,19]. In such a situation, not only the duration but also the cost and environmental impact of each activity depends on construction mode selection, which is in turn related to the assignment of construction materials and machines, crew formation, and overtime policies. In this case, therefore, a DTCETP-mP discrete mode selection for these variables is more practical than a continuous mode selection.

In the DTCETP-mP analysis, each subproject activity is assigned certain discrete well-defined construction methods so that the analysis can determine the best combination(s) of duration, cost, and environmental impact for all activities. The start time for each subproject activity is optimized by the flc-hGA, and all activities are executed in a certain order to achieve completion under budgetary, environmental, and cash flow constraints in any time period. The project completion date is as estimated by project experts and lies between the earliest start time for the first activity and the latest finish time for the last activity. The total project penalty costs, also assessed by these experts, comprise penalty costs for delays in all subprojects and for the environmental impact. Project planning involves the selection of proper methods, crew sizes, equipment, and technologies to ensure efficient activity completion.

Despite such careful planning, however, project completion inherently involves a trade-off between time, cost, and environmental impact. For example, using more productive resources or technologies (e.g., more efficient equipment, more workers, increased overtime) may save time but raise costs. Conversely, a reduction in time may lower environmental impact, while a reduction in cost may increase both time and environmental impact. These three important considerations of time, cost, and environmental impact are thus the criteria used to determine the best scheduling and combination for the DTCETP-mP, thereby pinpointing the optimal construction mode combination for all activities.

2.2 Assumptions

The DTCETP-mP model for the JHS-I makes the following assumptions:

(1) The DTCETP-mP comprises multiple projects, each containing several activities;
(2) The start time of each activity is dependent upon the completion of its predecessor;
(3) The capital used by all activities does not exceed the limited quantities in any time period, and the total project budget is within a predetermined limit;
(4) The environmental impact caused by the activities does not exceed the limited quantities in any time period, and the total project environmental impact is within a predetermined limit;
(5) When an activity begins, it cannot be interrupted;
(6) The managerial objective is to minimize the total project time, total tardiness penalty, and total environmental impact for all subprojects.

2.3 Model Formulation

The problem is represented on an activity-on-node (AON) network with a single starting and a single ending node, each of which corresponds to dummy activities. The analysis uses the following notation:

Index:
i : index of subproject in a project, where $i = 0, 1, 2, \cdots, I, S$ and T are dummy projects;
j : activity index in each subproject, $j = 0, 1, 2, \cdots, J$, and are dummy activity.

Parameters:
$\tilde{d}_{ij}$: processing time for activity j in subproject i;
$\tilde{t}_i^D$: specified project completion time for subproject i;
l_{ijc} : cost of activity j in subproject i per unit of time;
l_{ije} : environmental impact of activity j in subproject i per unit of time;
c_i^p : total penalty cost of subproject i per unit of time;
b_c : maximum cost allowable per unit of time for the subproject;
b_e : maximum environmental impact allowable per unit of time for the subproject;
B_c : maximum cost allowable per unit of time for the project;
B_e : maximum environmental impact allowable per unit of time for the project;
$\tilde{t}_{ij}^F$: finish time of activity j in subproject i;
V_{ij} : environmental impact of activity j in subproject i;
w_{ij} : weight of the environmental impact of activity j in subproject i;
S_p : set of activities in progress in period p
$S_p = \left\{ j \left| t_{ij}^S \leq p \leq t_{ij}^S + t_{ij}, i = 1, 2, \cdots I+1; j = 1, 2, \cdots J+1 \right. \right\}$;
$\Pr e(i)$: set of the immediate predecessors of subproject i;
$\Pr e(j)$: set of the immediate predecessors of activity j

Decision variables:
$\tilde{t}_{ij}^{S}$: start time for activity j in subproject i

Functions:
z_1 : total duration of the project;
z_2 : total penalty costs of the project;
z_3 : total environmental impact of the project

2.4 Multiobjective Model

The solution proposed by the multiobjective optimization model is based on managerial objectives and project constraints, which for clarity are discussed separately in the subsections below.
(1) Objective Functions

The first objective is to minimize total project time; that is, the sum of the completion times for all subprojects, which can be expressed as follows:

$$\min z_1 = \sum_{i=1}^{I} E(\tilde{t}_{iJ}^{F}). \tag{1}$$

The second objective is to measure and minimize the total cost by minimizing the total penalty costs of multiple projects:

$$\min z_2 = \sum_{i=1}^{I} c_i^p (E(\tilde{t}_{iJ}^{F}) - E(\tilde{t}_i^{D})). \tag{2}$$

The third and final objective is to minimize total environmental impact:

$$\min z_3 = \sum_{i=1}^{I} \frac{\sum_{j=1}^{J} V_{ij} \times w_{ij}(E(\tilde{t}_{iJ}^{F}) - E(\tilde{t}_i^{D}))}{E(\tilde{t}_{iJ}^{F})}. \tag{3}$$

(2) Constraints

Because a specific subproject must be completed before another subproject can be initiated (the precedence constraint of multiple projects), the model includes the following constraint:

$$E(\tilde{t}_{eJ}^{S}) \leq E(\tilde{t}_{i,1}^{S}) - E(\tilde{d}_{eJ}), \quad e \in \Pr e(i). \tag{4}$$

Likewise, because the start time of each activity is dependent upon the completion of some other activities (the precedence constraint of activities), the next activity must be started after a specific activity is completed:

$$E(\tilde{t}_{il}^{S}) \leq E(\tilde{t}_{ij}^{S}) - E(\tilde{d}_{il}), \quad l \in \Pr e(j). \tag{5}$$

The project is also subject to a limitation on the total capital and capital per time period,

$$\sum_{j \in S_p} l_{ijc} \leq b_c i = 1, 2, \cdots, I, \tag{6}$$

as well as on the total environmental impact and the environmental impact per time period:

$$\sum_{i \in S_p} \sum_{j \in S_p} l_{ijc} \leq B_c, \tag{7}$$

$$\sum_{j \in S_p} l_{ije} \times w_{ij} \leq b_e, \quad i = 1, 2, \cdots, I, \tag{8}$$

$$\sum_{i \in S_p} \sum_{j \in S_p} l_{ije} \times w_{ij} \leq B_e, \quad l_{ije} = \frac{V_{ij}}{d_{ij}}. \tag{9}$$

The nonnegative variables are described in the model by the following equation:

$$E(\tilde{t}_{ij}^S), E(\tilde{t}_{ij}^F), E(\tilde{d}_{ij}) \geq 0, \quad \forall i \in I, \quad \forall j \in J. \tag{10}$$

3 Hybrid Genetic Algorithm with Fuzzy Controller

Although mathematically, several Pareto optimal solutions are possible for the multiobjective model formulated above, in real-world construction, only one optimized solution is needed in each time-constrained decision-making situation. Hence, the multiobjective model is transformed into a single-objective model using a weighting method. Additionally, because accurately determining the GA parameters is especially important in solving large-scale problems like the JHS-I project, GA effectiveness is improved by adaptively regulating the crossover and mutation rate during the genetic search process using the fuzzy logic controller (flc) [16]. This regulation reduces CPU time and enhances optimization quality and stability by regulating the increasing and decreasing crossover and mutation rate ranges [1,5,13,14,20].

3.1 Overall Procedure for the Proposed Method

Solving the problem with the flc-hGA involves the following steps:

Step 1. Concentration of multiple objectives using the weight-sum procedure.
Step 2. Setting of the genetic algorithm parameters: population size, crossover rate, mutation rate, and maximum generation.
Step 3. Generation of an initial set of individuals.
Step 4. Choosing of the selection and hybrid genetic operators: crossover and mutation.
Step 5. Evaluation of the fitness value of the chromosome.
Step 6. Selection of the best total penalty for the minimized total project time and minimized environmental impact and storage of an alternative schedule for the minimized total project time.

Step 7. Check of the termination: If one individual has achieved the predefined fitness value, the process stops: otherwise, it goes on to step 8.

Step 8. Regulation of the mutation rate through adaptive use of the fuzzy logic controller; otherwise, the process returns to step 4.

The model uses two hybrid genetic operators, a position-based crossover and a swap mutation (SM) operator. The crossover operator randomly takes some genes from one parent and fills any vacuum with genes from the other parent by scanning from left to right, while the SM operator selects two projects at random and swaps their contents.

3.2 Hybrid Genetic Operators

The model uses two hybrid genetic operators, a position-based crossover and a swap mutation (SM) operator. The crossover operator randomly takes some genes from one parent and fills any vacuum with genes from the other parent by scanning from left to right, while the SM operator selects two projects at random and swaps their contents.

3.3 Fuzzy Logic Controller

The fuzzy logic controller (flc) is used to automatically tune the GA parameters whose determination is so important in large-scale problems. Here, only a mutation flc is used because in the proposed hGA, the effects of the crossover and crossover flc are almost the same. The main difference between the hGA and the flc-hGA is that a mutation flc based on the flc is implemented independently to adaptively regulate the mutation ratio during the genetic search process by considering the changes in average fitness in each parent and offspring population over two continuous generations. Mathematically, this procedure can be expressed as follows: letting $f(t)$ be the difference in the average fitness function between the t^{th} and $(t-1)^{th}$ generation, ε is a small positive number near to zero (in this paper, $\varepsilon = 0.1$), allowing the mutation ratio for the next generation to be derived using an if-then procedure:

(1) If $|f(t) - f(t-1)| < \varepsilon$, then the mutation ratio p_m for the next generation should be rapidly increased;
(2) If $f(t) - f(t-1) < \varepsilon$, then the mutation ratio p_m for the next generation should be decreased;
(3) If $f(t) - f(t-1) > \varepsilon$, then the mutation ratio p_m is selected for the next generation.

When $f(t)$ and $f(t-1)$ are the flc's inputs and the change in mutation ratio $m(t)$ is its outputs. $m(t)$ Once the input values are assigned, the scaling value $Z(a,b)$ can be determined by setting $\lambda \in [-1.0, 1.0]$ as the given values for regulating an increasing and decreasing range for the mutation ratio. The changes in the mutation ratios are then determined by $\Delta m(t) = \lambda Z(a,b)$ and the mutation ratio values for the next generation by $p_m(t+1) = p_m(t) + \Delta m(t)$, where $p_m(t)$ is the mutation ratio at generation t.

4 Case Study: The Jinping-I Hydropower Station

The Jinping-I (JHS-I) and Jinping-II Hydropower Stations, which epitomize the large-scale constructions of China's West-East Electric Transmission Project, have a combined capacity of 8,400 MW and were planned for Jingping River Bend, a major artery 150 Km in length whose downstream river section is separated from the opposite bank by only 16 km. Along that length, the elevation drops 310 m, creating an excellent site for hydroelectric production. Whereas JHS-I will rely on its high dam and reservoir to supply water, JHS-II, located 7.5 Km downstream (at coordinates $28°07'42''N101°14'27''E$) is diverted by a much smaller dam and will rely on the world's four largest and longest (16.7 km) diversion tunnels. This latter has a total installed capacity of 4,800 MW (8×600 MW), for a multiyear average annual generation of 24.23 TWh.

Because such scale and complexity presents computational challenges, our model includes only the elements typical of constructions, making it generalizable to similar projects. As previously emphasized, the study goal is to identify the most effective strategies and thus the most effective and viable project management methods for JHS-I, so that these may be implemented in future large-scale projects (Table 1).

Table 1. Subprojects and activities

A_1 Spillway project	A_{11} Earth-rock excavation A_{12} Concreting A_{13} Gates hoist equipment installation A_{14} Clearing up and finishing work
A_2 River diversion during construction	A_{21} Import and export hole dug A_{22} Concrete lining and check-gate installation A_{23} Lockup A_{24} Gen set installation A_{25} Substation construction and equipment installation
A_3 Dam construction	A_{31} Concrete cut-off wall A_{32} Dam foundation lock cut A_{33} Dam filled A_{34} Asphalt concrete watertight diaphragm
A_4 Power capacity of stream	A_{41} Diversion opening A_{42} Air pressure system A_{43} Second-stage cofferdam A_{44} Bore-hole A_{45} Concrete lining
A_5 Transport and power system	A_{51} Road clear A_{52} Warehouse and factory construction A_{53} Water supply A_{54} Power transmission project

4.1 Data Collection

The data for the JHS-I project, obtained primarily from the Ertan Hydropower Development Company, includes observations of managerial practice and interviews with designers, consultants, contractors, subcontractors, and a city government officer at the station. This data set is supplemented by information from prior research. The construction manager's project experience, in particular, was invaluable for researcher comprehension of the projects' specific nature and configuration. In addition to two dummy (start and end) projects, JHS-I has five subprojects: a transport and power system, river diversion during construction, dam construction, a stream power capacity project, and a spillway project (see Table 2).

Table 2. Detailed information for each activity

	Activity	DA	C	EI and W	PA	DP	PP
S	Dummy project			$S < 1, 2, 3$			
1	s (dummy activity)				s < 1,2	11	1<4,5
	1	2	2	20.0,0.05	1 < 3,4		
	2	5	2	32.6,0.055	2 < 4,t		
	3	5	2	16.7,0.03	3 < t		
	4	3	4	37.5,0.08	4 < t		
2	s (dummy activity)				s < 1,2	11	2 < 4,5
	1	4	2	23.5,0.037	1 < 3		
	2	2	3	24,2,0.048	2 < 3		
	3	4	2	23.8,0.038	3 < 4,5		
	4	3	1	19.8,0.053	4 < t		
	5	2	3	15.7,0.027	5 < t		
3	s (dummy activity)				s < 1	11	3 < 4,5
	1	2	3	25.0,0.027	1 < 2,3		
	2	5	3	23.6,0.019	2 < 4		
	3	3	3	21.3,0.033	3 < 4		
	4	3	1	21.6,0.041	4 < t		
4	s (dummy activity)				s < 1,2	11	4 < T
	1	2	1	11.6,0.027	1 < 3,4		
	2	5	2	16.3,0.047	2 < 4,5		
	3	4	3	11.4,0.028	3 < t		
	4	2	3	13.5,0.059	4 < t		
	5	4	1	11.5,0.037	5 < t		
5	s (dummy activity)				s < 1,2,3	11	5 < T
	1	1	1	21.6,0.031	1 < 4		
	2	2	2	26.7,0.0122	2 < 4		
	3	5	1	17.2,0.018	3 < 4		
	4	3	2	22.0,0.03	4 < t		
T	Dummy project						

Note: DA = expected value for activity duration (month), C = cost (million RMB), PA = activity predecessors, DP = expected value for project duration, PP = project predecessors, EI and W = environmental impact and weight.

Each activity must be performed in one of m_i possible modes, each with a corresponding duration, cost, environmental impact, and budget and cash flow limit (see Tables 3 and 4). Hence, each activity has a certain maximal cost and environmental impact unit requirement. Other relevant data for the analysis are as follows: The maximum capital (units: million RMB) and environmental impact for each time period are 14 and 10 units, $(B_c, B_e) = (12, 10)$, respectively, while the maximum capital (units: million RMB) and environmental impact for each subproject time period are 6 and 6 units, $(b_c, b_e) = (8, 6)$. The penalty cost for each subproject in each time period is $c_i^p = 12$ (units: million RMB). The evolutionary parameters are a population size of 20, maximal generation of 200, an optimistic-pessimistic index of $\lambda = 0.5$, and a weight for each objective of $\eta_1 = 0.5, \eta_2 = 0.2, \eta_3 = 0.3$. The remaining variables are activity duration, cost, activity predecessors, project duration, and project predecessors.

4.2 Case Study Results

To achieve managerial objectives, all project aspects must be optimized through the best possible arrangement of start times and construction modes for each subproject activity. According to the proposed model, for optimal scheduling, the JHS-I activities should be arranged in the order below and corresponding construction modes chosen to fulfil the decision maker's requirements:

$$S = \{P_2(1,2,3,5,4) \to P_3(1,2,3,4) \to P_1(1,3,2,4) \to P_4(2,1,3,5,4) \to P_5(2,1,3,4)\}.$$

Table 3. Optimal solution for the JHS-I case study

Optimal schedule	$S = \{P_2(1,2,3,5,4) \to P_3(1,2,3,4) \to P_1(1,3,2,4)$ $\to P_4(2,1,3,5,4) \to P_5(2,1,3,4)\}$
Sum of completion times for all subprojects	54 (month)
Total penalty cost	−18 (million RMB)
Total environmental impact	−6.88
Fitness value	0.871

Table 4. Contrast between the optimal solution and actual data

	z_1	z_2	z_3	Optimal order and executed mode
Actual project data	56	32	7.76	$S = \{P_1(2,1,4,3) \to P_2(2,1,3,4,5) \to P_3(1,3,2,4) \to P_4(2,5,1,4,3) \to P_5(3,2,1,4)\}$
Optimal result	54	−18	14.64	$S = \{P_2(1,2,3,5,4) \to P_3(1,2,3,4) \to P_1(1,3,2,4)$ $\to P_4(2,1,3,5,4) \to P_5(2,1,3,4)\}$
Net decrease in project value	2	50	15.5	
Rate of decrease	3.57%	156%	200%	

It should also be noted, however, that, because of cost and environmental impact constraints, certain noncritical activities in subproject 1 (2), executable between the 4^{th} and 9^{th} month, are flexible but other noncritical activities in subproject 3 (2), executable between the 3^{rd} and 8^{th} month, are not. The model thus defines the critical path meaningfully enough to be used in practical (rather than simply theoretical) scheduling. In particular, it enables the project manager to schedule activities according to the situation, which can be affected by such factors as available manpower, equipment, holidays and the need to harmonize with other parallel projects or activities. Moreover, even though a comparison of the actual project data with the optimal result generated by the flc-hGA reveals differences, these differences move in a positive direction. For example, the net decrease from plus to minus in each objective, when considered specifically for the changes in penalty costs and environmental impact, signals a shift from penalty to reward. Overall, the rate of this decrease for each objective, which ranges from 9.56% to 200%, signals an improvement in construction efficiency that could bring considerable economic benefit to any construction project, but especially to one that is large scale.

Admittedly, because the assumptions on which the mathematical model is based may generate certain modeling errors, the results do not represent a 100% optimal DTCETP-mP solution. Nevertheless, they can still serve as a useful optimal scheduling guideline for decision makers in current construction projects.

5 Conclusions and Implications for Future Research

The multiple objective optimization model developed in this paper extends a traditional single project model to an advanced multiple project trade-off model able to determine optimal scheduling and construction mode selection for the project at Jinping-I Hydropower Station. The model is designed to not only minimize duration and penalty costs but also reduce the DTCETP-mP's environmental impact. It thus controls for the project constraints of duration, cost, and environmental impact.

Because the DTCETP-mP is an NP-hard problem, the optimization method employs an flc-hGA to solve it. The analytical results for the JHS-I case study indicate that the trade-off between competing incommensurate objectives is actually dominated by the project manager's determination of the weight for each objective. Hence, decision makers can use the importance of project objectives to determine the optimum trends for time, cost, and environmental impact. One major advantage of the proposed method is that it enables decision makers to systematically and feasibly control the schedule according to an optimistic-pessimistic index. In addition, the flc-hGA developed for the study can be used to enhance optimization quality and stability.

The findings reported here suggest three important areas for future research, all of which warrant equal concern. Firstly, a more holistic measurement should be developed for environmental impact to ensure a more reasonable and effective model. Secondly, future studies should address more complex practical problems, such as those involving resource constraints, and uncertainties. Lastly, more efficient heuristic methods need to be developed to solve NP-hard problems with a greater number of constraints. The method and model proposed here offer a useful foundation for all such investigations.

Acknowledgements. This research was supported by the Chongqing advanced and applied basic research project (cstc2016jcyjA0442).

References

1. Afruzi EN, Najafi AA et al (2014) A multi-objective imperialist competitive algorithm for solving discrete time, cost and quality trade-off problems with mode-identity and resource-constrained situations. Comput Oper Res 50(10):80–96
2. Chen S, Chen B, Su M (2011) The cumulative effects of dam project on river ecosystem based on multi-scale ecological network analysis. Procedia Environ Sci 5:12–17
3. Eshtehardian E, Afshar A, Abbasnia R (2009) Fuzzy-based MOGA approach to stochastic time-cost trade-off problem. Autom Constr 18(5):692–701
4. Franck B, Neumann K, Schwindt C (2001) Truncated branch-and-bound, schedule-construction, and schedule-improvement procedures for resource-constrained project scheduling. OR Spectr 23(3):297–324
5. Gen M, Cheng R, Lin L (2008) Network models and optimization: multiobjective genetic algorithm approach
6. Harvey RT, Patterson JH (1979) An implicit enumeration algorithm for the time/cost tradeoff problem in project network analysis. Found Control Eng 4(3):107–117
7. Holland JH (1992) Adaptation in natural and artificial systems. MIT Press, Cambridge
8. Jeang A (2015) Project management for uncertainty with multiple objectives optimisation of time, cost and reliability. Int J Prod Res 53(5):1503–1526
9. Ke H, Ma J (2014) Modeling project time-cost trade-off in fuzzy random environment. Appl Soft Comput 19(2):80–85
10. Long LD, Ohsato A (2008) Fuzzy critical chain method for project scheduling under resource constraints and uncertainty. Int J Project Manage 26(6):688–698
11. Meier C, Yassine AA et al (2016) Optimizing time-cost trade-offs in product development projects with a multi-objective evolutionary algorithm. Res Eng Des 27(4):1–20
12. Monghasemi S, Nikoo MR et al (2014) A novel multi criteria decision making model for optimizing time-cost-quality trade-off problems in construction projects. Expert Syst Appl 42(6):3089–3104
13. Mungle S, Benyoucef L et al (2013) A fuzzy clustering-based genetic algorithm approach for time-cost-quality trade-off problems: a case study of highway construction project. Eng Appl Artif Intell 26(8):1953–1966
14. Said SS, Haouari M (2015) A hybrid simulation-optimization approach for the robust discrete time/cost trade-off problem. Appl Math Comput 259:628–636
15. Tareghian HR, Taheri SH (2007) A solution procedure for the discrete time, cost and quality tradeoff problem using electromagnetic scatter search. Appl Math Comput 190(2):1136–1145

16. Wang PT (1997) Speeding up the search process of genetic algorithm by fuzzy logic. In: Proceedings of 5th European congress on intelligent techniques and soft computing
17. Xu J, Zheng H et al (2012) Discrete time-cost-environment trade-off problem for large-scale construction systems with multiple modes under fuzzy uncertainty and its application to Jinping-II hydroelectric project. Int J Project Manage 30(8):950–966
18. Zheng H (2014) The fuzzy time-cost-quality-environment trade-off analysis of resource-constrained multi-mode construction systems for large-scale hydroelectric projects
19. Zheng H (2014) The fuzzy time-cost-quality-environment trade-off analysis of resource-constrained multi-mode construction systems for large-scale hydroelectric projects. Lect Notes Electr Eng 242:1403–1415
20. Zheng H, Zhong L (2017) Discrete time-cost-environment trade-off problem and its application to a large-scale construction project. Springer Singapore

Research on Equalization of Public Services Based on Changes of Urban and Rural Population Structure

Qian Fang(✉) and Yi Sheng

Sichuan Academy of Social Sciences,
Chengdu 610072, Sichuan, People's Republic of China
lily_lily009@163.com

Abstract. This paper starts the discussion from five aspects population distribution, age structure, employment characteristics, family characteristics and flow population, and conducts qualitative and quantitative analysis on the changes of urban and rural population of new period of China, so as to master the basic situation and development trend of urban and rural population. On this basis, the paper proposes suggestion and countermeasures on equalization of public services in urban and rural areas in our country, basically from three aspects-"increasing the flexibility of the urban and rural public service supply mode", "promoting the urban and rural service supply balance and accessibility" and "strengthening the effective supply to important groups and regions".

Keywords: Equalization of public services · Urban and rural areas · Population structure · Public service supply

1 Research Background

The problem of urban and rural areas is always a hot issue during the process of China's development. In November 2013, in the "Decision on Several Significant Problems on Comprehensively Deepening Reform by the Central Committee of the Communist Party of China" issued by the Third Plenary Session of the 18^{th} PCP Central Committee, it is clearly proposed to "propel the equalization of basic public services in urban and rural areas." There is no doubt that in the new period, the government has taken the perfection of urban and rural development integration mechanism as the main path for solving the problem of the unbalanced development between urban and rural areas, and has taken the equalization of public services as the main measure for relieving the conflict on development between urban and rural areas.

On the premise of the given orientation, it is urgent to conduct reasonable allocation on the public service resources of urban and rural areas and ensure the balance. In order to solve this problem, three aspects shall be taken into consideration: Firstly, what is the current situation of urban and rural population

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_145

structure? Secondly, how to define the equalization of urban and rural public services? Thirdly, how to realize the equalization of urban and rural public services? Domestic and foreign scholars have given some answers to the above questions, but not sufficient. Take the urban and rural population structure as an example. In 2011, the proportion of urban residents in our country exceeded that of the rural residents for the first time, which is called by media as “the reversal of urban and rural population structure since thousands of years in China” [3]; China has turned to the urban civilization stage with characteristics of industry and service industry from the giant agricultural country and agricultural civilization. It is noted in the “Social Blue Book” by Chinese Academy of Social Sciences that: “this is not only a change in the urban population percentage, but means the extremely profound change on the production method, occupational structure, consumption behavior, living method and value system of people” [12]. Is 2011 the time-point for the qualitative change of the urban and rural structure of China? If we narrow the scope to four major economic regions east region, middle region, west region and northeast region, we can find that such a so-called reversal has already appeared in Jiangsu, Zhejiang and Fujian as early as 2005. In the west region, the rural population in many regions is still more than that of the urban population (such as Guangxi, Sichuan, Guizhou, Yunnan, etc.), and no reversal has happened. The conclusion obtained from the average value of the statistical data is both scientific and some bit of biased, which shall not be taken too much consideration. Besides, the urban and rural population structure often discussed by some scholars is actually the urban and rural registered population structure, i.e., the personal status relationship of urban and rural residents, which is greatly different from the actual distribution of urban and rural population. In 2014, the population of separation of registered and actual residence in our country was 298 million, and the floating population was 253 million. It means that the era judging the population distribution according to family register has passed in our country. In brief, the urban and rural population structure in our country has come to a complex stage; from now on, equalization of public service will step into a period of overall promotion, accurate definition and courageous attempt. On the basis of accurate mastering of the change of urban and rural population structure, we will propose reform ideas and suggestion aiming at the new problem on urban and rural public service, which is an interesting topic, and also the original intension of this research.

2 Research Review

When conducting retrieval by taking CNKI as the retrieval source, “public service” and “public product” as the title and key word, and “urban and rural” as the title, more than 1800 results are obtained, in which there are 189 master and doctoral thesis, 149 conference papers, more than 570 papers on newspaper and 940 journal literatures.

There are 344 literatures with high quality, which are published on SCI, EI, Chinese core periodicals and CSSCI source periodicals. The following conclusion is obtained after conducting analysis on literatures by utilizing Citespace software.

Firstly, the high quality periodical achievements are concentrated during 2003–2016, mainly after 2010. Secondly, the top three authors in terms of the number of publishing are Wu Yemiao, Liu Chengkui and Yu Yaguai. No preferable cooperative relationship is formed between the authors, and most literatures are researched independently. Thirdly, the top three research institutions are the Institute of Public Administration of Sichuan University, the School of Government of Nanjing University and the Financial and Tax Institute of Shandong University of Finance. Fourthly, the research emphasis on urban and rural public service field in our country is concentrated in basic public service, equalization, urban and rural overall development, urban and rural gap, rural public products, sports public service and public finance. The research emphasis has been gradually transferred from the public product (2004), urban and rural overall development (2005), urban and rural public service (2006–2007) of the early stage to equalization of basic public service (2008–2010) and urban and rural basic public service (2010). Since "the 11th five-year plan", the domestic research has tended to specific contents such as social insurance, sports public service, infrastructure, etc. The research pays more attention to the evaluation on public service, to discuss on public service in combination of the theme of urbanization. Fifthly, there are two sudden nodes during the whole time period: public product and urban and rural overall development. The little node quantity indicates that the research on this field is not active enough, without significant emerging trend. Of course, it may be due to the short term of the research on this field in our country.

These researches mainly explore the equalization of public services from four aspects. First, it is the research on the specific contents of the basic public services. Secondly, it is the argument on the unbalanced situation of urban and rural public service. Just as what is proposed by Wu [14] that the unbalanced allocation of urban and rural public resources in our country is mainly expressed on infrastructure, basic education, social insurance and public medical health. According to Li [6], the unbalanced allocation of urban and rural public resource in our country is embodied on social insurance resource, social welfare resource, public health resource, basic education resource and infrastructure construction. Thirdly, explore the reasons for unbalanced urban and rural public service. It is regarded by some scholars that the subsidy on industry by agriculture and the subsidy on city by village in early stages of our country is the fundamental reason for unbalanced urban and rural public resource allocation [2]. It is considered by some scholars that the long term unbalanced urban and rural public resource allocation is related to the thoughts of "inevitable primary accumulation", "insufficient national financial resources" and "obstruction on efficiency by justice" in the society [15]; some scholars emphasize economic development level [5], and that the finance is the initiator of the evil [7]. Fourthly, explore the practice of promoting the equalization of urban and rural public service. Some scholars propose on ideology aspect to formulate the development strategy in accordance with the urban and rural integrated development, to increase the investment intensity in rural areas [18], to break through the functional restraint of "economic-oriented government", and to create public service-oriented government [16]; some scholars propose on financial aspect to take the investment by the Party Central Committee as the leading force, and to improve the efficiency of implementation of policy [11]. Some scholars think that it is needed to standardize the government expenditure behavior, and clarify the functions of various levels of governments, to increase the financial investment of the central and local governments [2]. Some scholars advocate to promote the balanced allocation of urban and rural public resources through marketization [1,19], but it is needed to avoid the corruption generated thereby. Some scholars think that it is available to conduct allocation on public resources by non-profit organizations [13]. Some scholars think that the balanced allocation of public resources is actually a problem about institutional arrangement in essence; To change the administrative philosophy of local government, to establish sustainable financial expenditure mechanism of rural public resource, to establish the peasants-participating public resource allocation decision mechanism, and to perfect the supervision and restraint mechanism of balanced allocation of social public resource, are four path choices for promoting the balanced allocation of public resources [6]; some scholars propose that it is necessary to realize the integration and association within the system, between the systems, between government and the society as well as among the diversified subjects of the society [4]. Besides, some scholars propose the method to improve the dual development on urban and rural areas from the aspect of specific field of public resources such as infrastructure, public education and medical health.

To sum up, the existing researches have laid firm foundation for us to continue the exploration on the urban and rural public service field, but there are still deficiencies on these researches. Firstly, compared with the discussion over the country (for example, 31 provinces and cities), there is obviously insufficient discussion on the equalization of urban and rural public services. Secondly, most researches are only conducted superficially on the unbalanced allocation of public service, resulting in superficial and impractical reform path. Thirdly, the existing researches are static, taking no population mobility and demand variability into consideration.

3 Changes on Urban and Rural Population Structure Characteristics in China

In recent ten years, there have been great changes on urban and rural population structure in China, which are embodied on aspects such as space, ages, family, education, income and employment. As for the population, the urban population in our country in 2011 exceeded more than 50% of the total population, forming a "neck and neck" situation on urban and rural population. On space distribution, the trend of migration from rural to urban area still goes on, and weight transition of urban and rural population and the adjustment of population density have become normalcy. On aspect of age groups, the fierce urban and rural population aging trend still goes on, and the situation of economic increase depending on "demographic dividend" has passed. On aspect of employment selection, the urban and rural population rushes to the tertiary industry (service industry), to make the employment transferred from "industry-oriented" to "service-oriented" form. With the resultant force of decreased family population and higher income, it has been the general situation for public service to lean to fields such as medical health and pension.

3.1 Changes on Space Distribution of Urban and Rural Population

Since the establishment of the nation, China has had an increasingly grown population, which increased to 1.36 billion in 2013 from 0.54 billion in 1949. The urban population has been on the rise, which increased to 0.73 billion in 2013 from 0.0577 billion in the early years of the new nation. Compared to the "up and down" situation of urban population, the rural population proportion was on a slow and continuous downtrend. In the early years of the new nation, the rural population in our country was near to 90%, which decreased to 46.27% in 2013.

Extend the analysis vision field to the four major economic regions, i.e., the east region, the middle region, the west region and the northeast region. The changes of urban and rural population show both generality and difference. On one hand, the changing trends of urban and rural population proportion in the east region, the middle region, the west region and the northeast region are almost the same, i.e., the proportion of urban population increases year-by-year, as shown in Table 1. On the other hand, there is significant difference on the proportion of urban population among the four economic regions. At the end of "the 10th five-year plan" (2005), the proportion of the urban population of the east region had been near to 60%, and the proportion of the urban population of the northeast had been near to 55%, while the proportion of the urban population

Table 1. Changes on proportion of urban population in the four major economic regions in China (2005–2013)

Region	2005	2006	2007	2008	2009	2010	2011	2012	2013
China	42.99	44.34	45.89	46.99	48.34	49.95	51.27	52.57	53.73
The east region	59.23	60.15	60.87	61.65	62.48	64.43	65.19	66.11	66.92
The middle region	37.58	38.96	40.27	41.73	43.03	44.44	46.28	47.98	49.26
The west region	35.18	36.15	37.31	38.53	39.61	41.45	42.81	44.26	45.43
The northeast region	54.77	55.15	55.42	56.22	56.39	57.04	57.98	58.75	59.35

Data source: Obtained by arrangement of "China Statistical Yearbook 2014" of the middle region and the west region was between 35%–38%. During 2006–2013, the urbanization of population in east region slowed down, with only 6.77% growth during 8 years; the middle region had the quickest growth speed, with 10.3% growth; the west region had a growth speed of 9.28%; the northeast region had the slowest growth speed, with only 4.2% growth during 8 years. Moreover, the active force for urbanization during the latest ten years in China was in the middle region and the west region, which are also the regions with the largest development potential in the future.

3.2 Changes on Age Structure of Urban and Rural Population

It is agreed by most domestic and foreign experts that "demographic dividend" is the main active force for the rapid development of 30 years since the reform and the opening-up policy. The young and middle-aged labor force plays a huge role in the rapid economic growing period. However, as time goes by, the population proportion of national labor force significantly decreases. During 2011–2014, the proportion of the population of 16–59 years' old decreased year-by-year, from 0.941 billion in 2011 to 0.916 billion in 2014. Correspondingly, the proportion of the population above 60 and 65 years' old increased year-by-year.

Up until the end of 2013, except for Tibet and Xinjiang regions, the 29 regions of the whole country had entered into the aging society to varying degrees. In which the aging degree of Tianjin, Shanghai, Jiangsu and Shandong of the east region, Anhui and Hunan of the middle region as well as Sichuan and Chongqing of the west region had exceeded than 10%. The top three regions of aging degree were Chongqing, Sichuan and Jiangsu.

Viewing from the four major economic regions, the proportion of the labor force population of 15–64 years' old in the east region is higher than that of the middle region and the west region, as shown in Table 2. In consideration that the permanent resident population of the east region occupies 37% of the 31 provinces and cities, which is much higher than that of the west region, the middle region and the northeast region, the development power in the east region is more sufficient and higher.

Seen from the aspect of dependency ratio, the national aging population dependency ratio increased from 8% in 1982 to 13% in 2013; the children dependency ratio decreased from 54.6% in 1982 to 22.2% in 2013. Aging of population

Table 2. Comparison on population age structure of the four major economic regions (2013)

Region	0–14(%)	15–64(%)	Over 65 (%)
The east region	14.87	75.52	9.61
The middle region	18.34	72.09	9.57
The west region	18.2	71.89	9.9
The northeast region	11.3	79.07	9.63

has a considerable influence on family and the society. Compared with the east region, the middle region and the northeast region, the aging population dependency ratio in the middle region is on the high side, and that in the west region and the northeast region is on the low side. The dependency ratio of children in the west region is the most highest, and that in the northeast region and the east region is on the low side. As a whole, the dependency ratio of families in the west region is the highest (38.29%), and that of the northeast region is the lowest (26.57%), and the middle region and the east region are between the two.

3.3 Changes on Employment Characteristics of Urban and Rural Population

During the 30 years since the reform and opening-up policy, the changes on employment structure of national urban and rural population are firstly embodied on employment population distribution of three industries; it appears in a situation that the proportion of the employment population of the primary industry decreases, and that of the secondary industry increases but with frequent fluctuation, and that of the tertiary industry continuously increases. The national total employment population in 2013 was 769.77 million, in which 382.4 million were urban employment population and 387.37 million were rural employment population. The employment population between urban and rural areas appeared in a "fifty-fifty" situation.

There are totally different employment structures in urban and rural areas. The urban employment population includes the people working in state-owned units, collective units, joint stock partnership units and joint ownership units, as well as in foreign-funded enterprises, private enterprises and individual enterprises. There are mainly three types of employment population in rural areas: the employment population in private enterprises, individual employment and faming population. Since 2004, the proportion of population working in private enterprises of rural employment population has exceeded than that of individual enterprises, showing the agriculture industrialization and enterprise-oriented development trend, as shown in Fig. 1.

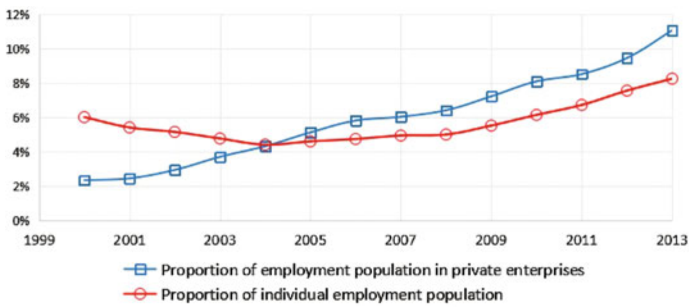


Fig. 1. Changes on proportion of employment population in private and individual enterprises in rural areas of China (2000–2013)

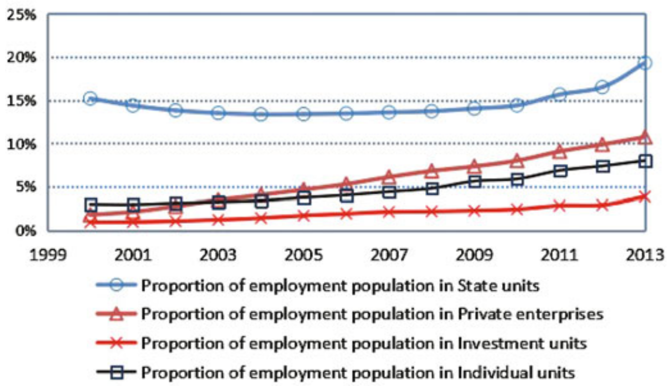


Fig. 2. Changes on proportion of urban employment population in China (2000–2013)

A large part of urban employment population "works in institutions as well as enterprise and public institutions of various levels", and please refer to Fig. 2 for the urban employment population situation. If the urban employment population working in state-owned enterprises, collective enterprises, joint-stock enterprises and associated enterprises is collectively called as "state enterprises", we can see two major changes of the national employment. Firstly, the employment population of "state units" appears in a slow rising trend. In the "state units", the employment population in state-owned units, collective units, joint stock partnership units and joint units decreases, and that in limited liability companies and joint-stock companies increases. These changes are corresponding to the state-owned enterprise reform and the separation of ownership and management, as well as the mixed ownership reform. Secondly, the employment population in private enterprises, investment units (foreign investment as well as investment made by businessmen of Hong Kong, Macau and Taiwan) and individual employment increases. The increase speed of the employment population of private enterprises is the fastest, followed by individual employment, and the investment unit is the slowest.

3.4 Changes on Family Characteristics of Urban and Rural Population

During 1949–2010, there was a small fluctuation in family scale in our country, decreasing from 4.43 persons for each household, which was the maximum value, to 3.10 persons for each household. In 2013, the national per household scale decreased to 2.98 persons for each household. When making comparison among the families in the east region, the middle region, the west region and the northeast region, the middle region has the largest family household scale, achieving 3.19 persons for each household. The northeast region has the smallest family household scale, with only 2.76 persons for each household. The middle region and the west region have the family household scales of 2.98 persons and 3.04 persons respectively.

From the comparison on urban and rural residents income situations among the east region, the middle region, the west region and the northeast region, we can know that the per capital income of the urban residents in the east region is the highest, and that in the west region is the lowest, and that in the middle region and the northeast region is between the two. The urban resident income in the east region is 1.44 times of that in the west region, and the per capita net income of rural residents is 1.76 times of that in the west region. Viewing from the data in the recent ten years, the proportion of income between residents in the east region and the west region gradually decreases, from 1.52 in 2005 to 1.43 in 2013, and the proportion of rural resident income decreases from 1.98 in 2005 to 1.76 in 2013.

According to the data, the top four expenditures for urban residents are "food", "transportation and communication", "culture, education and entertainment" and "clothing"; the top four expenditures for rural residents are "food", "residence", "transportation and communication" and "medical health". There are great difference on public service demands between rural residents and urban residents. "Transportation and communication" is the common part for urban and rural residents, but rural residents need more on the improvement of "residence" and "basic medical service", while urban residents pay more attention on "culture, education and entertainment".

It can be known from the comparison on urban residents expenditure in the east region, the middle region, the northeast region and the west region (as shown in Table 3) that the top four expenditures for urban residents in the four regions are "food", "transportation and communication", "culture, education and entertainment" and "residence". In which the consumption on "food" (37.07%) is the largest in the west region; the consumption on "transportation and communication" (16.36%) and "culture, education and entertainment" (13.35%) are the largest in the east region; the consumption on "residence" is the largest in the northeast region. It can be known according to the expenditure of rural residents among the east region, the middle region, the northeast region and the west region (as shown in Table 4) that the top four expenditures for rural residents in the four regions are "food", "residence", "transportation and communication" and "medical health". In which the consumption on "food" occupies the largest proportion (38.17%) in the west region; the consumption on "residence" (20.74%) occupies the largest proportion in the northeast region; the "transportation and communication" (13.22%) occupies the largest proportion in the east region; the consumption on "medical health" (12.16%) occupies the largest proportion in the middle region. Therefore, the urban and rural residents in the four regions have different emphasis on public service supply, and the quality and quantity of public service supply shall be improved according to the emphasis of resident expenditures.

Table 3. Proportion of per capita consumption expenditure of urban residents in the four major economic regions (2013)

Region	Unit	Food	Clothing	Residence	Household articles	Transportation equipment and communication	Culture, and entertainment	Medical education	Others health
National	%	35.02	10.55	9.68	6.74	15.19	12.73	6.2	3.88
The east region	%	35.33	9.04	9.45	6.61	16.36	13.15	5.78	4.29
The middle region	%	35.68	11.36	10.02	6.97	13.28	12.87	6.38	3.41
The north-east region	%	32.28	12.19	11.25	5.92	13.44	11.62	9.08	4.23
The west region	%	37.07	11.51	9.57	6.67	12.99	11.93	6.69	3.57

Data source: obtained by calculation and arrangement of "China Statistical Yearbook 2014"

Table 4. Per capital consumption expenditure of rural residents in the four major economic regions (2013)

Region	Unit	Food	Clothing	Residence	Household articles	Transportation equipment and communication	Culture, and entertainment	Medical education	Others health
National	%	37.67	6.62	18.62	5.84	12.01	7.33	9.27	2.64
The east region	%	37.35	6.66	16.98	5.66	13.22	8.12	8.86	3.16
The middle region	%	34.45	7.82	17.28	4.03	12.27	9.02	12.16	2.96
The northeast region	%	37.42	6.1	20.74	6.34	10.33	6.94	9.46	2.36
The west region	%	38.17	6.84	18.98	5.85	12.09	6.12	9.46	2.49

Data source: obtained by calculation and arrangement of "China Statistical Yearbook 2014"

3.5 Changes on Space and Characteristics of Floating Population

The floating population of China in 2014 was 0.253 billion, occupying 18.5% of the total population of the whole country. Liu, et al. [8] conducted measurement on the directivity and activeness of regional floating population in our country, and divide the floating population regional type into four types. In which the urban agglomeration regions such as the Pearl River Delta, the Yangtze River Delta and Beijing are the active regions for net inflow. The rural areas in the east region and the middle and west regions with concentrated population are the active regions for net outflow. The periphery or rural areas of the developed urban agglomeration in Zhejiang, Fujian, Guangdong in the east region are the active balanced regions. The vast middle region, the west region, Shandong, Hebei and north Jiangsu in the east region, Jilin and Liaoning in the northeast region are the non-active regions. There is relatively stable distribution proportion of floating population in the east region, the middle region and the west region during 2000–2010, with the increase rate of 115%–120%. The floating population in the east region occupies 2/3 of that of the whole country, and with rapid growth speed. The floating population concentration area in coastal regions gradually diffuses, with a continuous trend [9]. This trend is very significant in the Yangtze River Delta area, and the diffusion in the two floating population concentration areas of the Pearl River Delta and the Beijing-Tianjin-Hebei region is relatively limited. Megalopolises such as provincial capitals of inland regions attract a large amount of floating population; at the same time, there is significant northward movement on the distribution center of floating population (Table 5).

Table 5. Differentiation on zones of floating population in China [10]

Index	Year	The east region	The middle region	The west region
Total floating population (10 thousand persons)	2000	5110.4	1237.8	1552.6
	2010	10987.1	2661.9	3407.1
	The growth rate during the 10 years	115.00%	115.10%	119.40%
Proportion of floating population in the whole country (%)	2000	64.7	15.7	19.7
	2010	64.4	15.6	20
Proportion of floating family resister (%)	2000	11.1	3	4.4
	2010	22.1	5.8	8.8

When making comparison on the population floating regional difference in the east region, the middle region and the west region, we can find that the population floating in the east region is mainly the floating to other provinces, and that in the middle region is mainly the floating within county, and that in the west region is mainly the floating between provinces. As a whole, the floating between provinces occupies an important role; however, the within-province floating has a more fierce development trend, with the scale not less than the cross-province floating. Because of the lower cost and smaller obstruction, the floating within province of population will become the mainstream, and local urbanization will become the main method for the urbanization in our country.

Besides, the domestic population floating mode appears in significant "domestication trend". Most Chinese families adopt the "progressive floating mode", i.e., some family members enter into cities to work and live firstly. There are differences on the characteristics of family floating in different regions. For example, "the complete family type floating occupies the highest proportion in the middle region and the cross-county floating, which occupies a relatively low proportion in economic developed regions and under developed regions. This phenomenon illustrates that the threshold or living cost for family reunion in developed regions is relatively high, and the under developed regions fail to have an attraction for family reunion to floating population because of the lack of high quality resources" [17].

3.6 Countermeasures and Suggestions on Promoting of Equalization of Urban and Rural Public Service

Based on the change features on urban and rural population structure characteristics, the author proposes three suggestions and specific measures for promoting the equalization of urban and rural public services in China.

(1) Increase the Flexibility of Urban and Rural Public Service Supply Modes

Strengthen the statistical investigation, and accurately master the urban and rural population characteristics. It is suggested to constitute the yearly population sample survey working group by the statistical bureau and related departments by relying on existing population census organizations in our country, to develop yearly survey on population quantity, age, structure, distribution, residence and public service willing. Hold the joint conference of the statistic department, the public security department, the family planning department, the civil affair department, the environmental-protection department, the financial department and the population and society department according to the requirements on data collection, to clarify the rights and obligations of various departments on coordination for collection index. Establish urban and rural public service evaluation index system, to conduct analysis on data collected by each year by multiple forms including service outsourcing, cooperation between governments, franchising, subsidy and voluntary services, and the government releases the "Survey Report on Urban and Rural Public Services Situations in China" to the whole country each year. Change the method of allocating financial capital according to registered population, to promote the coordination between the urban permanent residence population and the central fiscal appropriation.

Eliminate the urban-oriented policies, and deepen the household registration system reform. Explore the way to take the ID card as the national use and unchanged for lifelong social guarantee number, and implement the social insurance one-card system for the national social insurance unities. On the premise of giving respect to the autonomous right of peasants to live and stay in cities, encourage them to go to the town. Prohibit taking the giving up of contractual right of land, right to the use of curtilage and the collective benefit distribution right as the conditions for peasants to settle down in cities. Broaden the residence migration policy, and take the legal stable residence or the legal stable employment registered residence as the basic conditions, to allow rural residents to apply for the registration of permanent residence household, and broaden the settling conditions for purchasing, building and inheriting of houses in cities, and simplify the population registration system with the principle of "conducting management on population through house, to achieving between population and household".

Promote the interaction of population, and construct integrated labor force market. It is suggested to drive the equalization of urban and rural public service by the urban and rural human resource integration. Establish the urban-rural unified employment standard, to desalt the localization and census register conditions of employment field, determine the subjective position of the labor force market, and realize the allocation of labor force resources by the market. Establish the labor force market with "online + offline" dual platforms, and improve the construction of information network, to establish the information system with the joint participation of enterprises, job seekers and occupation introduction institutions. Perfect the employee training and promotion system, and perfect the incentive mechanism, and encourage enterprises to maintain the employee training rights, and enlarge the promotion space for peasant workers, to ensure them to acquire the living capital in cities in an effective way.

(2) Promote the Balance and Accessibility of Urban and Rural Service Supply

Through giving directional difference to the rural area, relieve the unbalanced urban and rural structure. Further perfect the infrastructure construction of transportation in rural areas, to realize the smooth roads, and finish the construction of outward roads of poor villages. Improve the management and maintenance level of rural roads, and stable the road situation of rural roads, and extend the service life, and solidify the rural roads construction achievements. Promote the urban-rural integration, and perfect the urban system planning and village and town building plan for all the urban and rural areas. Pay more attention on aging population and school children on rural public service resources.

Promote development on small and middle-sized cities, and find out supporting point for city back seeding countryside. It is suggested to conduct comprehensive consideration on terrain and economic conditions on the basis of building the center cities in the middle region and the west region, to select a batch of type I and type small-sized cities, to cultivate and encourage them to grow larger and stronger, to form different layers of radiation circles, to provide substantial support for the development of rural areas.

Formulate differential equalized strategy according to regional characteristics. Conduct consideration by linking the infrastructure and the accessibility of public services with the space behaviors of residents, to strengthen the accessibility of public space, employment, administrative center and other comprehensive public services. Conduct different environment governance measures between urban and rural areas, and lay the emphasis on environment improvement (such as on aspects of drinking water, toilet, rubbish and waste water processing) in rural areas, and lay the emphasis on air quality and traffic congestion improvement in urban areas. According to the public service levels in the east regions, the middle region, the west region and the northeast region, formulate different propelling schemes. Highlight the characteristics of urban and rural areas, and strengthen the advantages of landscapes in urban and rural areas. Avoid rural areas copying the urban community mode, or substituting rural landscape with urban elements, and replacing greenbelts with firm grounds. It is not advocated to promoting the unified allocation standard between rural and urban areas, to avoid the loss of rural characteristics and folk culture in rural areas.

(3) Strengthen the Effective Supply to Important Groups and Key Areas

Improve the public service supply level and guide the peasant workers to integrate in cities in an orderly way. Perfect the infrastructure construction of the labor force market, and set up and integrate the "urban and rural labor force market information platform" integrated with enterprise employment information, labor force employment hunting information, industry employee information and employment policy. Increase the employment stability of peasant workers, and improve the vocational skill quality of peasant workers by means such as releasing free training tickets and implementing enterprise training fee deduction taxes and duties, to increase opportunities for employment. Increase the participation of the peasant worker group to urban public affairs, and set up peasant worker service platform in communities, to help them adapt to the city life as soon as possible by providing services such as formality agency, living guidance, employment consultation and children entrance information, etc.

Construct service network, so as to improve the medical and pension services for senior citizens in urban and rural areas. Conduct general investigation on the data, residence and family conditions of the senior citizens in national urban and rural areas, to establish documents for the senior citizens. Arrange medical health institutions or nursing institutions for the aged in urban and rural areas, to build the pension community integrated with medical and pension services, to form the urban-rural covering and reasonably functional medical and pension service networks with sharing resources between pension service and medical health service. Improve the ability of door-to-door service for senior citizens by the basic level medical health institutions, and encourage medical institutions to provide green channel for senior citizens to seek for medical service, and encourage nursing institutions for the aged to provide medical service for senior citizens. Guide the enterprises to develop integrated pension communities integrated with residence, health care, pension, recovery and entertainment; guide the community health service institutions to conduct transformation and upgrading, and encourage the comprehensive medical institutions to cooperate with nursing institutions for the aged, to give play to the radiation and leading functions of medical alliance.

References

1. Chen J (2008) Market-oriented allocation of public resources: theoretical basis, risk analysis and route selection. J Municipal Party Committee Party Sch Yinchuan 1:59–62 (in Chinese)
2. Chen J (2010) Empirical analysis on different influences of fiscal expenditure behavior on infrastructures of urban and rural areas. Xinjiang Acad Soc Sci 3:24–29 (in Chinese)
3. Daily H (2011) Reversal of urban and rural population structure. Hubei Daily
4. Jing T (2015) Base line justice: welfare basis point of balanced development between justice and development. J Beijing Technol Univ (Soc Sci Ed) 1:1–7 (in Chinese)
5. Lei X, Zhang N (2012) Research overview on urban and rural overall development of social insurance. Stud Soc Secur 1:91–96 (in Chinese)

6. Li J, Zhou Q (2015) Supervision mechanism and route selection on balanced allocation of urban and rural public resources. Shanghai City Manage 1:34–37 (in Chinese)
7. Li Y, Peng Y (2016) Research on fiscal policy of the equal access to basic public services between urban and rural areas. Finan Theory Pract 3:56–61 (in Chinese)
8. Liu S, Deng Y, Hu Z (2010) Division method and space distribution characteristics for regional types of floating population in China. J Geogr Sci 10:1187–1197 (in Chinese)
9. Liu T, Qin Y, Cao G (2015) Space pattern evolution mechanism of floating population in China and urbanization effect-analysis based on population census county basis of 2000 and 2010. J Geogr Sci 4:567–581 (in Chinese)
10. Liu Y, Qi Y, Cao GJ (2015) Evaluation mechanism of space pattern of floating population in China and urbanization effect-analysis based on the county data of population census of 2000 and 2010. J Geogr Sci 4:2–13 (in Chinese)
11. Luo Y (2010) Analysis on economic effect of difference on infrastructure of urban and rural areas in China-based on space panel measurement model. Chin Rural Econ 3:60–72 (in Chinese)
12. Social Sciences Academic Press (2011) Social blue book: Chinese social situation analysis and prediction. Social Sciences Academic Press
13. Qin B, Xiao S (2013) Public resource allocation on non-profit orgnizations. Finan Econ 2:41–45 (in Chinese)
14. Wu L, Xu C (2014) Discussion and analysis on balanced allocation on urban and rural public resources. Acad Commun 5:131–134 (in Chinese)
15. Wu Z (2008) The chinese society stepping to justice. Shandong People's Publishing House, Jinan
16. Wu Z (2013) The institutional constraint and solution on urban and rural education integration. J South China Normal Univ (Soc Sci Ed) 1:29–32 (in Chinese)
17. Yang J, Chen C (2013) Current state and characteristics of floating population family-oriented pattern: analysis on characteristics of floating process. Popul Dev 3:2–13 (in Chinese)
18. Zheng Y, Liu Z, Wang Y (2012) The infrastructure construction of urban and rural overall development under the new rural background. Soc Sci Res 6:25–29 (in Chinese)
19. Zhou C (2013) Strategic thinking on promotion of public resource allocation market-oriented reform. Forum People 11:66–67 (in Chinese)

Economic Cycle, Accounting Conservatism and Financial Constraints

Hong Wang[1(✉)], Pan Liang[1], and Juqiu Deng[2]

[1] Business School, Sichuan University, Chengdu 610065, People's Republic of China
979166074@qq.com

[2] Economics School of Sichuan University, Chengdu 610065, People's Republic of China

Abstract. This paper adopts cashflow sensitivity of cash holding model to investigate the influence of economic cycles and accounting conservatism on corporate financial constraints, and to investigate the different impact of the two mechanisms from the property right perspective. Our results show that the recession of economic cycles aggravates the degree of corporate financial constraints, and this negative influence is more apparent on non-SOEs than SOEs. Moreover, the increase of accounting conservatism can release corporate financial constraints, and the alleviating effect of accounting conservatism is more significant on non-SOEs.

Keywords: Economic cycle · Accounting conservatism · Financial constraints

1 Introduction

Financial constraints can effect the corporate innovation [9] and investment efficiency [10], which seriously restrict the further development of corporates and ultimately affect economic growth. The World Bank report [7] shows that 75% of non-financial listed companies in China choose financial constraints as a major obstacle to business development, the highest percentage in 80 surveyed countries. Chinese corporates are facing the test of financial constraints.

China's economy characteristics has been markedly different from the past 30 years. It mainly expresses in the following aspects: (i) economic growth tailed off to around 7%, or lower; (ii) a large sum of money flows into the real estate and infrastructure construction instead of the real economy; (iii) the real economy such as manufacturing achieve consistently poor performance. The lack of capital in the real economy is detrimental to the long-term sustainable economic growth. And capital flowing into the real estate sector may not be able to promote economic growth and add to economic stability. Instead, it may boost the housing prices and make the real economy more difficult. Against this background, it is of practical significance to study the causes and solutions of the financial constraints of traditional enterprises.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_146

Theories have studied the causes of financial constraints from information asymmetry, agency costs, firm size and political connection [6,19,23]. However, all these studies are carried out from the microperspective. After taking these factors into account, the phenomenon that the degree of financial constraints of Chinese corporates fluctuates with the macroeconomic can't be explained. Trying to find out the reason, this article raises the first question: since the fluctuation of economic cycle is one of the main characteristics of market economy, do the economic cycles affect the corporate financial constraints? If so, what is the influence mechanism?

Accounting conservatism is an important index to measure the quality of accounting information [3,4]. Fair value has been questioned after 2008 financial crisis, while theories begin to rethink the positive effect of accounting conservatism [17,24]. The increases in conservatism improve the firm information environment and lead to subsequent decreases in information asymmetries between firm insiders and outsiders [15]. And accounting conservatism reduces under investment in the presence of information frictions [2]. So this article raises the second question: is it easier for the enterprises with high accounting conservatism to get financing in the capital markets? Can accounting conservatism ease the financial constraints, especially in the economic recession?

Using the financial data of Chinese listed A - share listed manufacturing companies, we test whether and how economic cycle and accounting conservatism affect corporate financial constraints to obtain the empirical evidence. Our first contribution is our explicit focus on the impact of economic cycles on corporate financial constraints. Typically, existing studies tend to focus on studying the cause of financial constraints from the enterprise's perspective. Macroeconomics and enterprise behaviors are disjointed in existing research [11]. There is a strong evidence, however, that the economic cycles can influence corporate financial constraints. As financial constraints may be specific to different property rights, we investigate the different impact from the property right perspective. It can broaden the research horizon on influencing factors of financial constraints from purely concerning about the internal factors of corporates to the external macroeconomic environment.

The second main contribution of this paper is that we pertain to providing an approach to ease financial constraints suffered by Chinese traditional enterprises. We suggest that it's beneficial for corporates to concern about the external business environment, especially the economic cycles. As for the impact of economic cycles-there are two representative views: the first is that the effect of the economic cycles on each firm is exactly the same, so changes in the economic cycles do not lead to additional losses for individual firm [16]; Another view is that the profitability of a firm is linked to the economic cycles, and its investment and business strategies are effected by economic cycles [18]. Our research confirms that the recession of the economic cycle aggravates corporate financial constraints. Enterprises should take advance measures to cope with the lough business environment. Moreover, we study the positive effect of accounting conservatism on decreasing corporate financial constraints and make a useful supplement to the economic consequences of accounting conservatism. It will inspire the traditional enterprises to improve accounting information quality for decreasing financial constraints.

The remainder of the paper is organized as follows. Section 2 explains the relevant theoretical analyses and our hypotheses. Section 3 presents our research sample and model. Section 4 shows our main empirical findings, and Sect. 5 presents our conclusion.

2 Theoretical Analyses and Hypothesis

In the economic recession, reduced external demand leads to production downsizing, which inevitably damage enterprises profitability. Further, there is a reduction in internal funds. So, enterprises must rely on more external financing to meet funding needs. Bernanke and Gertler [6] point out that enterprises suffer higher external financing costs and more serious cash-flow problems during the downturn. This is mainly caused by the tough external business environment. Besides that, the deterioration of corporate operating situation leads to the increase of information asymmetry. At the same time, investors become more cautious with big confidence attack. The stock market is in depression, so enterprises are difficult to raise enough funds by issuing shares. In addition, financial institutions such as banks may reduce the loan amount and raise the loan interest rate to control the default risk, increasing the difficulty of obtaining loans from financial institutions. According to financial accelerator theory, enterprises suffer more serious financial constraints because the external capital suppliers require a higher risk premium to compensate for possible loss risk in a recession. On the contrary, the optimistic expectations help to ease the financial constraints in the macroeconomic boom. It can be seen that there is a significant negative correlation between the degree of financial constraints and the economic cycles. Therefore, this paper gives the first hypothesis:

Hypothesis 1: The economic cycles can influence the degree of corporate financial constraints. During the boom, the degree of financial constraints is relatively low, while in the economic recession, the degree of financial constraints is high.

Based on the soft budget constraint theory, the influence of economic cycle on financial constraints can be different due to property rights. On the one, SOEs take more political tasks than non-SOEs. The government is more likely to give stealth or explicit support to help SOEs out of troubles when SOEs involve in financial risks [21]. On the other, the long-term mutually beneficial contract between the state-owned banks and the SOEs can alleviate the information asymmetry between the two. So SOEs enjoy preferential treatment in bank loans. The Property rights, to a certain extent, can weaken the influence of economic cycle on corporate financial constraints. So, this paper gives the second hypothesis:

Hypothesis 2: Compared with SOEs, non-SOEs suffer more serious financial constraints.

Enterprises with poor financial status may carry out earnings management to meet the financing requirements to get funds. Creditors and shareholders are in a weak situation with less information and can't effectively supervise the use of funds. So enterprises are more prone to moral hazard and adverse selection. Creditors and shareholders will require a higher risk premium for taking a greater default risk. Therefore, under the same conditions, the higher degree of information asymmetry, the more serious financial constraints enterprises suffer. Some studies show that accounting conservatism can help to ease the agency conflict. Beaver and Ryan [5] find that accounting conservatism mitigates agency conflicts between management and creditors by undervaluing net assets and discouraging management from transferring creditor interests to shareholders through asset substitution. Lafond and Roychowdhury [14] argue that management share holding ratio is negatively correlated with accounting conservatism because accounting conservatism eases agency conflicts between management and shareholders by suppressing earnings management and over-investment. As a signal transmission mechanism, accounting conservatism can effectively reduce the agency costs and avoid the moral hazard and adverse selection caused by information asymmetry. It shows that accounting conservatism plays a positive role in easing financial constraints. Based on this, *Hypothesis 3* is proposed.

Hypothesis 3: Under the same conditions, the improvement of accounting conservatism can alleviate the corporate financial constraints.

Due to the special national situation of our country, the SOEs still occupy the dominant position in the market economy. SOEs not only have more resources and policy advantages, but also enjoy the implicit guarantees from government. When the SOEs fall into financial troubles, the government is more willing to provide financial assistance and banks are more willing to provide loans. It reduces investors' concern of accounting conservatism, that is, accounting conservatism has little effect on alleviating the financial constraints of SOEs. On the contrary, accounting information quality of non-SOEs is one of the most important factors considered by the external capital suppliers. The conservative accounting policies can help enterprises obtain financing from banks and investors. In the downturn, non-SOEs suffer more serious financial constraints, so it is extremely important for non-SOEs to obtain financing from external markets. Using accounting conservatism to show their stable profitability and good operating conditions can alleviate the financial constraints. Viewed from the bank point, SOEs also have the credit quotas granted by banks. So the financial constraints suffered by SOEs aren't serious even in the downtown and the financial constraints caused by economic downturn will be more serious in non-SOEs. Based on the above analysis, the hypothesis 4 is proposed.

Hypothesis 4: Under the same conditions, the increase of accounting conservatism weakens the negative influence of economic recession on corporate financial constraints, and the interaction is more significant in non-SOEs.

3 Research Sample and Model Design

3.1 Definition of Variables

(1) Financial constraints

Both cash flow sensitivity of investing model and cash flow sensitivity of cash holding model are widely used in present research about financial constraints. However, due to the completely opposite conclusions in different studies using cash flow sensitivity of investing model, this paper uses cash flow sensitivity of cash holding model to measure financial constraints. And Almeida et al. [1] also prove that the cash flow sensitivity of cash holding model can overcome the defect of cash flow sensitivity of investing model.

(2) Accounting conservatism

Accounting conservatism is typically defined in the empirical accounting literature as the differential verifiability required for recognizing gains as compared to losses [4, 22]. More precisely, conservatism is the practice of reducing earnings in response to "bad news", but not increasing earnings in response to "good news". Basu [4] uses the positive stock returns to express good news and negative stock returns to show bad news, and construct the following model to measure enterprises' accounting conservatism.

$$\mathrm{EPS}_{i,t}/P_{i,t-1} = \alpha_0 + \alpha_1 \mathrm{EPS}_{i,t} + \alpha_2 \mathrm{RET}_{i,t} + \alpha_3 DR_{i,t} \times \mathrm{RET}_{i,t} + \varepsilon_{i,t}, \tag{1}$$

where i indicates the firm, $\mathrm{EPS}_{i,t}$ is earnings, $\mathrm{RET}_{i,t}$ is returns, and $DR_{i,t}$ is a dummy variable that equals 1 if $\mathrm{RET}_{i,t}$ is less than 0, and 0 otherwise. So α_2 is the good news timeliness measure and α_3 is the incremental timeliness for bad news over good news, or conservatism.

Khan and Watts [13] improved the Basumodel. They considered firm-year specific coefficients α_2 (timeliness of good news) and α_3 (conservatism) can be expressed by linear functions of firm-year characteristics that are correlated with the timeliness of good news and conservatism:

$$\text{G-SCORE} = \alpha_2 = \mu_1 + \mu_2 \mathrm{SIZE}_{it} + \mu_3 \mathrm{LEV}_{it} + \mu_4 MB_{it}, \tag{2}$$

$$\text{C-SCORE} = \alpha_3 = \lambda_1 + \lambda_2 \mathrm{SIZE}_{i,t} + \lambda_3 \mathrm{LEV}_{i,t} + \lambda_4 MB_{i,t}, \tag{3}$$

where SIZE_{it} is the natural log of the market value, MB_{it} is the market-to-book ratio, and LEV_{it} is the debt-to-equity ratio. Replacing α_2 and α_3 in Eq. (1) by Eqs. (2) and (3), respectively, yields the following empirical regression model:

$$\begin{aligned} \mathrm{EPS}_{i,t}/P_{i,t1} = {} & \alpha_1 + \alpha_2 DR_{i,t} + (\mu_1 + \mu_2 \mathrm{SIZE}_{i,t} + \mu_3 \mathrm{LEV}_{i,t} + \mu_4 MB_{i,t}) \\ & \times \mathrm{RET}_{i,t} + (\lambda_1 + \lambda_2 \mathrm{SIZE}_{i,t} + \lambda_3 \mathrm{LEV}_{i,t} + \lambda_4 MB_{i,t}) \\ & \times DR_{i,t} \times \mathrm{RET}_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{4}$$

(3) Economic cycle

Domestic scholars use different methods to divide the economic cycle. Jiang and Liu [12] used GDP growth rate to distinguish different economic cycle. Shi and Zhang [20] divided the economic cycle based on the relationship between the real economic growth rate and the potential economic growth rate. In addition, Dong Jin (2006) [8] used a variety of measurement methods such as linear trend method, H-P filter method, B-P filter method and the production function method to divide China's economic cycle. This paper references the research of Jiang and Liu [11], etc., and use the annual GDP growth rate as the surrogate variable of the economic cycle (Clycle_GDP). First, GDP growth rate as a continuous variable can effectively avoid the error of using subjective dummy variables; And the use of continuous variable can also effectively avoid the systemic impact of institutional change, such as the implementation of new accounting standards, the split share structure reform.

(4) Property rights

This paper divides enterprises into SOEs and non-SOEs only, and researches the influence of economic cycle and accounting conservatism on financial constraints from these two respects. The property rights are defined as dummy variables, which are judged according to the ultimate controller of the listed company. If it's a SOE, it equals 0 and 1 otherwise.

(5) Control variables

Referencing the model of the existing research, this paper selects the operating cash flow (Cfo), enterprise size (Size), asset liability ratio (Lev), growth (growth), profitability (Roa), short-term borrowing difference (△Debt) as the control variables. The more sufficient operating cash flow, the more liquidity the company has, indicating that the company is in a good condition. So the expected sign is negative. Company size is defined as the natural logarithm of the total assets at the end of the year. Large scale companies are more resistant to risks. So, the default risk is smaller and the expected sign is positive. The higher the debt-to-asset ratio, the weaker the solvency, so the expected sign is positive. Enterprises will undertake a higher cost of capital if they have a higher growth and more likely to fall into financial distress. But on the other hand, they have better development potential and less defaultrisk. Profitability is the enterprises' operating efficiency. The better the enterprise's profitability, the less likely the default occurs. The sign is expected to be positive.

3.2 Model Design

As for the research model of financial constraints, Almeida's cash flow sensitivity of cash holding model has been recognized by many scholars. The model suggests that if enterprises suffer higher financial constraints, more cash will be held from the enterprise's cash flows to prepare for investment opportunities, and the cash flow sensitivity of the enterprise will be higher. In order to verify the four hypotheses, this paper builds extension models based on Almeida's cash flow sensitivity of cash holding model.

(1) Test the influence of economic cycles on corporate financial constraints

In order to test the influence of the economic cycle on financial constraints suffered by enterprises, this paper adds a cross-term between the operating cash flow and the economic cycle to cash flow sensitivity of cash holding model, and prove whether the economic cycle affect the financial constraints by checking the coefficients of the cross terms. As for the *Hypothesis 2*, the group multiple regression analysis is performed according to the different property rights.

$$\begin{aligned}\triangle\text{Cash} = &\beta 0 + \beta 1\, Cfo + \beta 2\,\text{Cycle\_GDP} + \beta 3\,\text{Cycle\_GDP} \times \text{Cfo} + \beta 4\,\text{Size}\\ &+ \beta 5\text{Lev} + \beta 6\,\text{Roa} + \beta 7\,\text{Growth} + \beta 8\,\triangle\text{Debt} + \varepsilon\end{aligned} \tag{5}$$

(2) Test the influence of accounting conservatism on corporate financial constraints

In order to test the influence of accounting conservatism on the degree of financial constraint, this paper adds the cross-term between accounting conservatism and operating cash flow to cash flow sensitivity of cash holding model, and confirm the role of accounting conservatism by observing the coefficients of the cross terms. As for the *Hypothesis 3*, the group multiple regression analysis is performed by property rights.

$$\begin{aligned}\triangle\text{Cash} = &\beta 0 + \beta 1\,\text{Cfo} + \beta 2\,\text{Conserv} + \beta 3\,\text{Conserv} \times \text{Cfo} + \beta 4\,\text{Size}\\ &+ \beta 5\,\text{Lev} + \beta 6\,\text{Roa} + \beta 7\,\text{Growth} + \beta 8\,\triangle Debt + \varepsilon.\end{aligned} \tag{6}$$

(3) Test the influence of economic cycle and accounting conservatism on corporate financial constraints

In order to test the interaction of economic cycles and accounting conservatism on the degree of financial constraints, this paper adds the cross-terms of economic cycle, operating cash flow and accounting conservatism to cash flow sensitivity of cash holding model. We observe the coefficients of the three cross-terms to verify how the financial cycles and the accounting conservatism interact the corporate financial constraints. And the group multiple regression analysis is performed by property rights also.

$$\begin{aligned}\triangle\text{Cash} = &\beta_0 + \beta_1\,\text{Cfo} + \beta_2\,\text{Conserv} + \beta_3\,\text{Cycle\_GDP} \times \text{Cfo}\\ &+ \beta_3\,\text{Cycle\_GDP} \times \text{Cfo} \times \text{Conserv} + \beta_3\,\text{Conserv} \times \text{Cfo}\\ &+ \beta_4\,\text{Size} + \beta_5\,\text{Lev} + \beta_6\,\text{Roa} + \beta_7\,\text{Growth} + \beta_8\,\triangle\text{Debt} + \varepsilon.\end{aligned} \tag{7}$$

3.3 Sample Selection and Data Sources

The sample consists of all A-share listed Chinese manufacturing enterprises between 2007 and 2015. The sample selection criteria are as follows: (1) Excluding enterprises with missing data; (2) Excluding companies with negative net assets; (3) excluding enterprises with uncertain property rights. Finally, 6933 samples are obtained, including 3975 SOEs and 2958 non-SOEs. In addition, extreme values of 1% and 99% of the relevant variables are Winsorized to eliminate the effects of extreme values. All financial data comes from CSMAR database and we use Stata11.0 to analysis the data (Table 1).

Table 1. Sample distribution

Year	2008	2009	2010	2011	2012	2013	2014	2015	Total
Number	593	626	675	853	989	1066	1066	1065	6933

4 Empirical Results

4.1 Descriptive Statistics

Table 2 provides descriptive statistics for the variables, while the whole sample is grouped by property rights. Table 3 reports the descriptive statistics of the variables in SOEs and non-SOEs. It can be seen that the mean value of △Cash is −0.00048, the minimum value is −0.759 and the maximum value is 0.636, which indicates that enterprises have large differences in cash holdings; And the mean value of △Cash of non-SOEs is significantly larger than that of SOEs, indicating that non-SOEs are more cautious than SOEs; The mean of Cfo is 0.0472, the minimum value is −4.27 and the maximum is 0.549, indicating that cash flows of different enterprises is also a gap. The mean of Conserve is 0.0397, the minimum is −0.00223 and the maximum is 0.276, which shows that accounting conservatism are widely used in Chinese enterprises, but there are big differences among enterprises. The mean of Conserving SOEs is 0.0397, and the mean of the non-SOEs is 0.0476, which indicates that non-SOEs have higher accounting conservatism than SOEs. The mean of Lev is 0.432, indicating that debt financing is one of the

Table 2. Descriptive statistics of all samples

Variable	N	Mean	Std	Max	Min	Med
△Cash	6933	−0.00048	0.0935	0.636	−0.759	0.000434
Cfo	6933	0.0472	0.0919	0.549	−4.27	0.0452
Cycle_GDP	6933	0.0841	0.0129	0.104	0.0690	0.0770
Conserv	6933	0.0397	0.0203	0.276	−0.00220	0.0395
Size	6933	21.88	1.194	26.96	17.47	21.74
Lev	6933	0.432	0.213	2.992	0.00708	0.431
Roa	6933	0.0404	0.0873	4.837	−1.292	0.0353
Growth	6933	0.376	9.396	665.5	−1	0.104
△Debt	6933	0.0108	0.0785	0.566	−3.04	0.000485

Table 3. Descriptive statistics of variables grouped by property rights

Variable	N	Mean	Std	Max	Min	Med	Mean T test
SOEs							
△Cash	3975	−0.0107	0.102	0.548	−0.732	−0.007	10.6247***
Cfo	3975	0.05	0.078	0.549	−0.656	0.0477	−2.9576***
Cycle_GDP	3975	0.0826	0.0126	0.104	0.069	0.077	11.4272***
Conserv	3975	0.0338	0.0189	0.276	−0.0015	0.033	29.7379***
Size	3975	21.58	0.983	25.88	17.47	21.49	25.7232***
Lev	3975	0.371	0.199	2.992	0.0071	0.362	29.4408***
Roa	3975	0.0484	0.0629	0.429	−1.164	0.0442	−8.8354***
Growth	3975	0.445	10.94	665.5	−1	0.121	−0.7081***
△Debt	3975	0.0135	0.0701	0.566	−0.787	0	−3.3414***
Non-SOEs							
△Cash	2958	0.0132	0.0781	0.636	−0.759	0.0088	10.6247***
Cfo	2958	0.0434	0.108	0.484	−4.27	0.0403	−2.9576***
Cycle_GDP	2958	0.0861	0.013	0.104	0.069	0.077	11.4272***
Conserv	2958	0.0476	0.0193	0.175	−0.0022	0.0491	29.7379***
Size	2958	22.29	1.325	26.96	18.37	22.09	25.7232***
Lev	2958	0.514	0.202	1.867	0.0224	0.529	29.4408***
Roa	2958	0.0298	0.111	4.837	−1.292	0.0248	−8.8354***
Growth	2958	0.283	6.793	363.1	−0.957	0.0831	−0.7081***
△Debt	2958	0.0071	0.0885	0.351	−3.039	0.0013	−3.3414***

t statistics in parentheses, $^{*}p< 0.1$, $^{**}p< 0.05$, $^{***}p< 0.01$

important financing channels. The mean of the growth is 0.376, the minimum is −1, and the maximum is 665.5. Assets scale and debt level of SOEs are higher than those of non-SOEs, but the non-SOEs have higher growth and profitability.

4.2 Correlation Analysis

Table 4 reflects the results of the correlation analysis. It can be seen that the relationship between the dependent variable and independent variables is consistent with the hypothesis. Cfo is positively correlated with △Cash, indicating that A-listed manufacturing enterprises generally have financial constraints. Cfo is negatively correlated with Cycle_GDP, which indicates that the economic cycle has obvious inhibitory effect on the external financing ability of enterprises. Conserv is negatively correlated with Cfo, which indicates that the improvement of accounting conservatism can alleviate the financial constraints suffered by enterprises; In the following empirical regression analysis, a more comprehensive summary will be made. In addition, all the correlation coefficients are less than 0.6, indicating that there is no multicollinearity between them.

Table 4. Correlation analysis results of major variables

Variable	ΔCash	Cfo	Cycle_GDP	Conserv	Size	Lev	Roa	Growth	ΔDebt	State
ΔCash	1.000									
Cfo	0.135***	1.000								
Cycle_GDP	0.021***	−0.007***	1.000							
Conserv	0.164***	−0.127***	0.098***	1.000						
Size	0.161***	0.081***	−0.107***	0.415***	1.000					
Lev	0.165***	−0.129***	0.096***	0.999***	0.402***	1.000				
Roa	0.117***	−0.140***	0.053***	−0.323***	0.010	−0.319***	1.000			
Growth	0.036***	−0.006	−0.005	0.024*	0.021*	0.024	0.010	1.000		
ΔDebt	0.070***	0.053***	0.056***	0.124***	0.082***	0.121***	−0.330***	0.059***	1.000	
State	−0.127***	0.036***	−0.136***	−0.336***	−0.295***	−0.333***	0.106***	0.009	0.041***	1.000

t statistics in parentheses, $^{*}p< 0.1$, $^{**}p< 0.05$, $^{***}p< 0.01$

Table 5. Regression analysis of economic cycles and financial constraints

Independent variable	All samples	SOEs	non-SOEs
	Regression model (5)		
Cfo	0.692*** (4.412)	0.200*** (0.764)	0.947*** (5.417)
Cycle_GDP	0.020*** (6.640)	0.009** (2.149)	0.025*** (5.986)
Cycle_GDP×Cfo	−2.253*** (−5.006)	−1.910*** (−3.069)	−2.254*** (−3.575)
Size	0.003*** (2.933)	0.002 (1.419)	0.007*** (3.684)
Lev	0.076*** (12.554)	0.020*** (2.642)	0.121*** (12.622)
Roa	0.249*** (17.808)	0.215*** (14.116)	0.282*** (9.995)
Growth	0.000* (1.888)	0.000** (2.209)	0.000 (0.857)
ΔDebt	0.139*** (9.171)	0.096*** (5.071)	0.162*** (6.978)
State	−0.017*** (−7.294)		
_cons	−0.105*** (−4.777)	−0.043* (−1.747)	−0.217*** (−5.939)
N	6933	2958	3975
r2	0.093	0.078	0.100
r2_a	0.092	0.075	0.098
F	78.804	31.183	54.939

t statistics in parentheses, $^{*}p< 0.1$, $^{**}p< 0.05$, $^{***}p< 0.01$

Table 6. Regression analysis of accounting conservatism and financial constraints

Independent variable	All samples	SOEs	non-SOEs
	Regression model (6)		
Cfo	0.396***	0.280***	0.510***
	−13.309	−6.297	−12.609
Conserv	−3.088** (−2.431)	−1.819 (−0.929)	−4.774*** (−2.880)
Conserv× Cfo	−4.458*** (−7.303)	−2.913*** (−3.483)	−4.162*** (−4.560)
Size	0.003***	0.001	0.009***
	−2.598	−0.784	−5.004
Lev	0.403***	0.225	0.570***
	−3.357	−1.218	−3.649
Roa	0.248***	0.252***	0.070**
	−16.884	−14.918	−2.379
Growth	0.000*	0.000**	0
	−1.955	−2.169	−0.594
△Debt	0.149***	0.100***	0.262***
	−10.333	−5.446	−11.399
State	−0.016*** (−7.007)		
_cons	−0.127*** (−5.399)	−0.053** (−1.966)	−0.281*** (−7.379)
N	6933	2958	3975
r2	0.125	0.096	0.152
r2_a	0.124	0.093	0.15
F	109.992	39.098	89.005

t statistics in parentheses
*p< 0.1, **p< 0.05, ***p< 0.01

4.3 Empirical Results and Analysis

(1) Economic cycle and financial constraints

The regression results of the regression model (5) are shown in Table 5. The regression coefficients of Cfoare positive at the significance level of 1%, whether in the regression of all samples or the sample of SOEs and non-SOEs, indicating that financial constraints is a serious problem in both SOEs and non- SOEs. And the coefficient of Cfo in non-SOEs is 0.947, which is significantly larger than 0.200 in SOEs, indicating that the financial constraints suffered by SOEs are more significantly. Then the coefficients of the cycle-GDP $*$ Cfo are analyzed. The coefficients are negative in the three regression at the 0.01significantly level, which shows that the economic cycle has a s negative correlation with the corporate

Table 7. Regression analysis of economic cycles, accounting conservatism and financial constraints

Independent variable	All samples	SOEs	non-SOEs
	Regression model (7)		
Cfo	0.318***	0.162***	0.465***
	−11.472	−3.823	−12.289
Cycle−GDP	0.021***	0.011**	0.022***
	−6.642	−2.464	−5.015
Conserv	−0.388 (−0.291)	−1.511 (−0.753)	−1.228 (−0.689)
Conserv× Cfo	−1.192*** (−2.845)	−0.659 (−1.101)	−1.713*** (−3.037)
Cycle-GDP∗ Cfo∗ Conserv	−31.434*** (−4.684)	−5.870* (−0.626)	−35.547*** (−3.651)
Size	0	0	0.005***
	−0.351	−0.194	−2.811
Lev	0.149	0.191	0.239
	−1.181	−1.011	−1.429
Roa	0.283***	0.268***	0.109***
	−19.459	−15.717	−3.676
Growth	0.000**	0.000**	0
	−2.02	−2.179	−0.654
△Debt	0.155***	0.102***	0.272***
	−10.773	−5.504	−11.839
State	−0.018*** (−7.741)		
_cons	−0.075*** (−3.078)	−0.035 (−1.275)	−0.205*** (−5.122)
N	6933	2958	3975
r2	0.129	0.095	0.159
r2_a	0.128	0.092	0.157
F	93.165	30.905	75.076

t statistics in parentheses: $^{*}p< 0.1$, $^{**}p< 0.05$, $^{***}p< 0.01$

financial constraints. The first Hypothesis has been confirmed. The coefficient of Cycle-GDP ∗ Cfo of SOEs is −1.910, while the coefficient of Cycle-GDP ∗ Cfo of non-SOEs is −2.254, which shows that the financial constraints of non-SOEs are more affected by the changes in the economic cycle. The Hypothesis 2 has been confirmed.

(2) Accounting Conservatism and Financing Constraints

The regression results of the regression model 6 are shown in Table 6. It should be noted that the coefficients of the cross-variable Conserv×Cfo are significantly negative in the three regression results, indicating that the increase of the accounting conservatism can alleviates the financial constraints suffered by the enterprises. The coefficient in SOEs is −2.913 and the coefficient in non-SOEs is −4.162, which shows that the degree of accounting conservatism of non-SOEs has a greater influence on financial constraints, the hypothesis 3 has been confirmed.

(3) Economic cycles, accounting conservatism and financial constraints

The regression results of model (7) are shown in Table 7. The coefficients of the cycle-GDP × Cfo × Conserv are significantly negative in the three regression results, indicating that the improvement of the accounting conservatism weakens the negative influence of economic cycles. In other words, enterprises with higher accounting conservatism can get the external financing easier during the downturn. And the coefficient of SOEs is −5.870, while the coefficient of non-SOEs is −35.547, which shows that the positive role of accounting conservatism in easing financial constraints during the downturn can be more obvious in non-SOEs. The *Hypothesis 4* has been confirmed.

5 Conclusion

This paper examines the interaction of the economic cycle and accounting conservatism on corporate financial constraints using the data of A-share listed manufacturing companies from 2008 to 2015 from the property right perspective. It is found that the macroeconomic recession aggravates the financing difficulty of enterprises and cause more serious corporate financial constraints. And this negative influence is more significant on non-SOEs. Besides, the improvement of corporate accounting conservatism can help to alleviate the financial constraints suffered by enterprises, and this positive influence is more significant on non-SOEs as well. We test the dynamic response of enterprises in the face of the economic cycle recession by combining the economic cycle with the micro-enterprise behavior. The micro-enterprisescan weaken the negative influence of the economic recession by increasing accounting conservatism, which not only helps to understand micro-transmission mechanism of economics cycle, but also helps enterprises to improve cash using efficiency.

The shortcomings of the study: As the national macro-data is annual announcement, only the annual data of the sample companies can be chosen in order to match with the macro data. It reduces the length of research period and may not accurately reflect the influence of macroeconomic fluctuations monthly on corporate financial constraints. In addition, the accurate measurement model is the basis of empirical research. Although there are many measurement methods abroad, there is no conclusive conclusion. Domestic research in this area is a reference to foreign research. Therefore, the focus of future research is create a more accurate accounting conservatism measurement method based on China's institutional background.

Acknowledgements. This study is founded by Sichuan University (skqy201523) and Social Science Annual Project of Sichuan Province (SC16TJ011). We are very grateful for their support.

References

1. Almeida H, Campello M, Weisbach MS (2004) The cash flow sensitivity of cash. J Finance 59(4):1777–1804
2. Balakrishnan K, Watts R, Zuo L (2016) The effect of accounting conservatism on corporate investment during the global financial crisis. J Bus Finance Account 43(5–6):513–542
3. Ball R, Shivakumar L (2004) Earnings quality in uk private firms: comparative loss recognition timeliness. J Account Econ 39(1):83–128
4. Basu S (2013) The conservatism principle and the asymmetric timeliness of earnings. Contemp Account Res 24(1):3–37
5. Beaver WH, Ryan SG (2005) Conditional and unconditional conservatism: concepts and modeling. Rev Acc Stud 10(2):269–309
6. Bernanke B, Gertler M (1989) Agency costs, net worth, and business fluctuations. Am Econ Rev 79(79):14–31
7. Claessens S, Tzioumis K (2006) Measuring firms' access to finance. World Bank
8. Dong J (2006) Measuring China's business cycles. Econ Res 7:41–48 (in Chinese)
9. Efthyvoulou G, Vahter P (2016) Financial constraints, innovation performance and sectoral disaggregation. Manchester Sch 84(2):125–158
10. Guariglia A, Yang J (2015) A balancing act: managing financial constraints and agency costs to minimize investment inefficiency in the Chinese market. J Corp Finance 36:111–130
11. Jiang GA, Rao P (2011) Macroeconomic policies and corporate behaviorłłbroaden accounting and corporate finance research horizon. Account Res 3:9–18 (in Chinese)
12. Jiang L, Liu X (2011) Economic cycle fluctuations and the cash holding behavior of listed companies. Account Res 9:40–46 (in Chinese)
13. Khan M, Watts RL (2009) Estimation and empirical properties of a firm-year measure of accounting conservatism. J Account Econ 48(2–3):132–150
14. Lafond R, Roychowdhury S (2008) Managerial ownership and accounting conservatism. J Account Res 46(1):101–135
15. Lara JMG, Osma BG, Penalva F (2014) Information consequences of accounting conservatism. Eur Account Rev 23(2):173–198
16. Lucas RE (1978) Asset prices in an exchange economy. Econometrica 46(46):1429–1445
17. Macve RH (2015) Fair value vs conservatism? aspects of the history ofaccounting, auditing, business and finance from ancientmesopotamia to modern China. SSRN Electron J 47(2):124–141
18. Gregory MN (2009) Principle of economics. Tsinghua University Press, Beijing
19. Shen CH, Lin CY (2015) Political connections, financial constraints, and corporate investment. Rev Quant Finance Account 47(2):1–26

20. Shi X, Zhang S (2010) Behavior of substitution between trade and bank borrowing through economics cycles: evidence from China. J Manage Sci China 13(12):100–122 (in Chinese)
21. Sun ZA, Liu F, Li ZC (2005) Market development, government influence and corporate debt maturity structure. Econ Res J 5:52–63
22. Watts RL (2003) Conservatism in accounting part I: explanations and implications. Soc Sci Electron Publ 17(3):207–221
23. Weiss A (1984) Information imperfection in the capital market and macroeconomic fluctuation. Am Econ Rev 74(2):194–199
24. Wier HA (2009) Fair value or conservatism: the case of the gold industry. Contemp Account Res 26(4):1207–1233

Traffic Lights Dynamic Timing Algorithm Based on Reinforcement Learning

Chenqing Lu[1], Feng Wen[1(✉)], and Mitsuo Gen[2]

[1] School of Information Science and Engineering, Shenyang Ligong University, Shenyang, The People's Republic of China
wenfeng_mail@163.com

[2] Fuzzy Logic Systems Institute, Tokyo University of Science, Tokyo, Japan

Abstract. Traffic light control is an issue that metropolitan cities are confronted with because of ever-increasing vehicles growth. A proper timing strategy can release the congestion level and minimize vehicle delays better than standard fixed signal timing (FST) strategy. In this paper, a self-adaptive Q-learning based traffic light green time control strategy (QTGCS) is designed to reduce the congestion. Q-Learning acts as the learning mechanism for light at each intersection as an independent agent to release itself by operating traffic light duration dynamically. Q-Learning is based on discrete representation of state and action spaces. In order to make QTGCS not independent of the environment and designer criterion, Fuzzy Q-learning based traffic light green time control strategy (FQTGCS) is proposed by applying the fuzzy logic control to optimize timing strategy based on the number of vehicles receiving from the adjacent intersections. We show through simulation that our algorithm effectively over the real road networks.

Keywords: Light control · Reinforcement learning · Q-learning · Fuzzy logic control

1 Introduction

Traffic congestion has been an urgent problem of metropolitan cities in the world. It is generally recognized that traffic signal improvements offer the biggest payoff for reducing congestion and increasing the effective capacity of existing road networks, and the adaptive traffic signal control systems hold the most promise for improvement. Reinforcement learning approach implicitly models the dynamics of complex systems by learning the control actions and the resulted changes of traffic flow. Meanwhile, it seeks the (sub)optimal signal plan from the learned input-output pairs. RL, usually formalized as a framework of Markov decision process (MDP), assumed that an intersection behaves similar to an intelligent agent learning to plan green times in each cycle using current traffic information.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_147

Specifically, Q-learning obtains Q-function by learning an action-value function. Thus, Q-learning is well-suited RL algorithm that obtains an optimal action-selection policy for given traffic state in any road networks. If traffic light indicates a suitable signal for vehicles, unnecessary waiting time of road users and the level of congestion state will be reduced compared to standard fixed signal timing (FST) strategy.

There are lots of researches aimed at reducing the congestion by the road users through RL methods. The isolated traffic signal was defined as an agent of RL, agent try to learn relationships between actions and their effect on the environment [1]. Extensive studies tried to implement RL or integrate with other algorithms for light timing, such as neuro-fuzzy [8] and kernel method to approximate Q-function, which is applied to estimate the feature vector [6]. Multi-objective RL proposed for traffic signal control in [7], in this approach, consolidating all rewards in one Q-function to obtain a global objective instead a multi-objective optimization problem. In [5], Q-learning algorithm acts as learning mechanism that react with environment for light intersections to release itself and in [4,9] taking advantage of fuzzy logic, the state and actions is set by a fuzzy algorithm which can be learned from an isolated intersection environment. In consideration of the correlation with adjacent intersections, there are also few researches aimed to coordinate light timing strategy together in small area separately [2], but how to form specific cooperation region not mentioned especially in large-scale road networks.

Fuzzy logic has the advantage of interpreting the linguistic values in terms of logical variables. With the aid of this, the traffic congestion situation can be clearly classified and comprehended [3]. Most of the fuzzy models regularly query the traffic conditions in order to decide whether to extend or terminate a current green phase based on Q value or reward value but not congestion relation on the adjacent sections. In this paper, we formulate the traffic light timing problem as an MDP, design a self-adaptive Q-learning based traffic light green time control strategy(QTGCS). In order to make QTGCS not independent of the environment and designer criterion, Fuzzy Q-learning based traffic light green time control strategy (FQTGCS) is proposed based on the number of vehicles receiving from the adjacent intersections. The indistinguishable boundary can be used evaluate the state approached next step. In the development process, we use SUMO simulator, an open-source software to help the real-world traffic light simulation. We show that FQTGCS yields a policy that has better performance than both the FST and QTGCS.

2 Traffic Lights Timing Strategy

2.1 Traffic Phase Model

The traffic light infrastructure consists of roads and nodes. A road connects two nodes, and can have several lanes in each direction. Phase is the permission which determines the duration of green time for traffic flows in one or some directions at the intersections. It plays an important role in ITS especially in light timing method. In this paper, we have defined two phase models: one for three-arm intersection phase model is shown in Fig. 1.

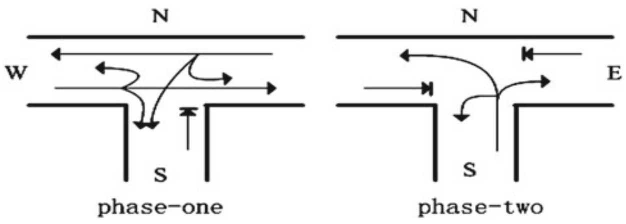


Fig. 1. Three-arm intersection phase model

Another four-arm intersection phase model is shown in Fig. 2. We can see easily from the figure that each model including four-phase consist of left-turning, right-turning, straight-going, and U-turn.

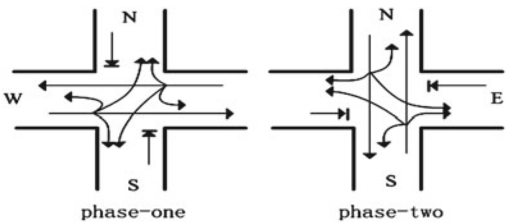


Fig. 2. Four-arm intersection phase model

In Fig. 1, phase-one denotes for the east-west direction, traffic light turns out to be green at a given time and vehicles get the pass permission, and for the south direction, the light is red at an three-arm intersection; phase-two is quite the reverse at that junction, the south direction for green light, and for the east-west direction, the light is red. In Fig. 2, phase-one denotes light turns green in the east-west direction and vehicles get the pass permission, and for the south-north direction, the light is red at a four-arm intersection; phase-two also denotes at that junction on the opposite way. In a FST strategy, the phases are set in a round-robin (RR) manner and the duration of green time for each phase is assigned in advance.

Due to the dynamic and uncertain nature of the traffic environment, we suppose that the phase of release order is changing frequently for light timing. The study will set phase selection strategy according to vehicles density. If the vehicles density in phase-one direction is higher, then choose phase-one; otherwise choose phase-two. In such way, which phase obtained the evacuation right can be easily determined.

$$\text{phase}(t) = \begin{cases} \text{phase-one} & \rho_{\text{phase-one}} > \rho_{\text{phase-two}} \\ \text{phase-two} & \rho_{\text{phase-one}} \leq \rho_{\text{phase-two}}, \end{cases} \tag{1}$$

where, phase(t) is the phase we have choose by phase selection strategy, $\rho_{\text{phase-one}}$ is vehicles density of phase-one direction, $\rho_{\text{phase-two}}$ is the density of phase-two direction.

2.2 MDP Framework

An MDP provide a coherent theoretical framework for which action is chosen in state s at each time step, and finally find a sequential decision by controller. In traffic light timing, we regard each traffic light as an control agent, and describe the process as a four-tuple (S, A, P, R), where S and A are the state space and action space, respectively, and every element $s \in S$ is called a state. Every element $a \in A$is called an action; $P : S \times A \times S$ is the transition probability which is a finite set of probability distributions in which an element $p(s, a, s')$ denotes the probability given the current state-action pair (s, a) to next state s'; and $R : S \times A \times S$ is the reward function, and $r(s, a, s')$ denotes immediate control performance received after transitioning from state s to state s', due to action a, where s,$s' \in S$, and $a \in A$. Hence, the whole traffic environment state is Markov, and the goal is to choose the optimal policy π^* that maximizes the total discounted reward, as:

$$\pi^* = \arg\max_{\pi} V^{\pi}(s). \tag{2}$$

In Eq. (2), $V^{\pi}(s)$ is the value function accumulated under the policy, which defined as follows:

$$V^{\pi}(s) = \sum\nolimits_{s' \in S} P(s, a, s')[R(s, a, s') + \gamma V(s')], \tag{3}$$

where $\gamma \in (0, 1)$ is the discount factor, we use the value function denote the discounted sum of the rewards.

In reinforcement learning, we can easily formulate light timing strategy within the MDP framework described, and the light at a signalized intersection as a learning agent. With this, light agent observe the congestion situation of the environment around, and take an action which based on exploration strategy, act on the environment and obtain a reward to evaluate the action. The light agent learning through the trial-error experience to achieve the optimal policy.

2.3 Q-Learning Algorithm

MDPs required having a good model in its implementation. Therefore, in this research we use Q-learning which from reinforcement learning with a model-free approach, and the traffic light at junctions as a learning agent. The light agent are not necessary to know how the environment works. It works by estimating the values of state-action pair, called Q-value, which represents the maximum discounted sum of future rewards an agent except to receive if it starts in state s, choose action a and then continues to follow an optimal policy.

This paper designed a distributed traffic light control strategy based on Q-learning. In such traffic light timing control model, each traffic light is an agent in Q-learning, agent chooses green time of traffic light as the action of agent, the density of vehicles in traffic lane at the intersection as the state, vehicles' average traveling time as the reward of Q-learning in traffic lane. Learning system interact with the environment constantly to get feedback and adjust map strategy of state to action.

The update of Q-learning formula is as follows:

$$Q(s,a) = (1-a)Q(s,a) + a\left[R(s,a) + \gamma Q^{\max}(s',a')\right], \tag{4}$$

where, α is the learning rate, and is the discount factor to decide convergence rate, $R(s,a)$ in function is the reward incurred by the light agent for taking an action a when the current state s and end up in state s', $Q(s,a)$ is the sum of future rewards to evaluate the performance of action we have took. It is an iterative method that modifies the estimates of Q-value. In the learning process, according to the Q-learning update rule set so that, the estimates converge to the optimal Q-value and the light agent's knowledge tends to be precise.

In this paper, Q-values are updated by the Q-learning update rule using Boltzmann based exploration strategy. The light agent decides its action based on the current learnt Q-value, the next section with Boltzmann exploration strategy that is used by the light agent to update Q-value and subsequently for choosing actions, the formulate is defined as follows:

$$p[a\,|s] = \frac{e^{\frac{Q(s,a)}{-t}}}{\sum_{a \in A} e^{\frac{Q(s,a)}{-t}}}, \tag{5}$$

where A is action space and τ is temperature parameters, $p[a\,|s]$ is probability that select a action in state s. The states represent the traffic density in the lanes connected to the junction described in Sect. 2.4, the action which represents the time for the highest congestion phase get the green light on.
Light timing control algorithm process based on Q-learning is shown below:

Step 1. initialize Q-learning algorithm parameters;
Step 2. analyzing the state of surrounding environment;
Step 3. according to the state, the light agent will select by action selection strategy;
Step 4. action act on the environment, and make environment changed, meanwhile the light agent obtain the reward according to reward function, and update the Q-value by this reward;
Step 5. if not satisfy the stop condition, return to step 2.

2.4 Fuzzy Logic Control Optimization Strategy

Consider the incoming traffic flows react on the congestion level of joint junction, and even if calculate the optimal state-action pair, there must be some disparities between actual optimal solution. For the consideration above, in this section, we propose a method to optimize action solution based on fuzzy logic control strategy. With fuzzy logic strategy, the learning agents coordinate with joint agents by information sharing, namely, the light agent allot green time to the phase that maintained the highest congestion level and obtain the status of approach phase flows by the vehicle information, these data are used to decide whether to extend or shrink the current light through-phase to achieve action solution optimization point. Furthermore, we use the density of vehicles to measure phase congestion level, and defined model discussed above in Fig. 3.

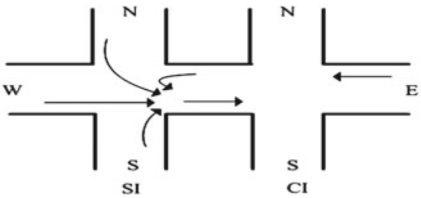


Fig. 3. Phase flow model

In this method, a fuzzy description of the traffic density during each phase at the intersection and its joint intersection, and the reward are implemented respectively, means fuzzify. After that, we develop fuzzy rulebase and initialize the reward values by the if-then relations.

(1) Fuzzy sets

A fuzzy set is a pair (U, A) where for each element $x \in U$, here U is a set, and A, a membership function, is the degree of membership of x in (U, A). The closer $A(x)$ to 1, the higher likelihood of $x \in A$, and the closer $A(x)$ to 0, the lower would be. We use trigonometric membership function in this paper. The domain of traffic density during the phase at the intersection is $[0, 1]$, the density level will be divided into four categories, density level below 0.25, density level between 0.25–0.5, density level 0.5–0.75, and the density level above 0.75, furthermore the fuzzy sub-sets of traffic density is "zero" (Z), "low" (L), "medium" (M), "high" (H), the membership function of each sub-sets is shown in Fig. 4:

The domain of the action reward is the real unit interval $[-2, 2]$, the four fuzzy sub-sets of reward is "zero" (Z), "small (S), "medium" (M), "medium big" (MB), "big" (B), the membership function of subsets is shown in Fig. 5:

Traffic light optimize action selection strategy according to density situation at current intersection (CI) and its joint intersection (JI), to maintain synergy between traffic lights in a small scale, and improve the traffic capacity of the total traffic road network efficiently.

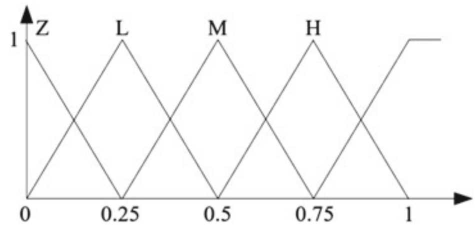


Fig. 4. Membership function for the phase of vehicle density

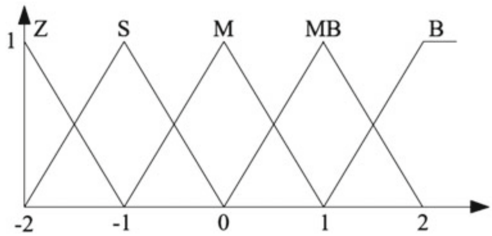


Fig. 5. Membership function for the reward

(2) Fuzzy database

In this section, we develop the fuzzy database as follows:

- If the traffic density of the phase at current intersection is M, at the same time, the approach density of traffic flows at joint intersection is L, then we can define the action reward is M;
- If the traffic density of the phase at current intersection is H, the approach density of traffic flows at joint intersection is Z at the same time, then the action reward is MB.

Thus with linguistic descriptions of traffic density as input, so the phase of vehicle density is determined by the fuzzy rules given in Table 1 described as:

$$\begin{aligned} &\text{If } CI \text{ is } TD_i \text{ and } JI \text{ is } TD_j, \text{ then } T_{\text{green}} \text{ is } r, \\ &(i, j = Z, L, M, H; r = Z, S, M, MB, B) \end{aligned} \tag{6}$$

where CI is the current intersection and its phase density TD_i, TD_j is the density approach from the joint intersection JI, r describes the green time to be extend or shrink. The fuzzy database is defined in Table 1

3 Experiments

In order to verify the validity and correctness of the traffic light control strategy. In this paper, the vehicle induction system (SVIS) is based on the shortest path

Table 1. Fuzzy database

Action reward		Corresponding intersection phase density			
		Z	L	M	H
Current intersection phase vehicle density	Z	Z	S	M	MB
	L	S	S	M	MB
	M	S	M	MB	MB
	H	M	MB	MB	B

algorithm. Simulation and experiment on the road network through the open-source software SUMO simulator, which is wrote by java, the road network is a part of the U.S. State of Vermont. In the formula about traffic light control strategy is based on Q-learning, we set α is 0.7, and γ is 0.9. The road network is shown in Fig. 6 (Table 2):

Table 2. Road network information

Attribute	Quantity
The number of traffic lights (3-way or 4-way intersection)	51
Number of road sections	206
Starting point	8
Terminal point	8
Departure speed (v/h)	1440
Total simulation time	15000

The data of the algorithm is evaluated in this experiment is:

(1) Count the quantity of vehicles in traffic system per 100 s;
(2) Count intersection vehicles average travel time in each phase.

Fig. 6. Road network

In SVIS, in traffic light control strategy based on Q-learning algorithm (QTGCS), light timing based on action information, and in the fixed-time green light control strategy (FTGCS), traffic light phase update based on fixed-time 50 s set by system. The evaluation data of SVIS and QTGCS were compared with the evaluation data obtained by SVIS and FTGCS. The quantity of vehicles in traffic system is shown in Fig. 7, and intersection vehicles average travel time in the lane is shown in Fig. 8 as follows:

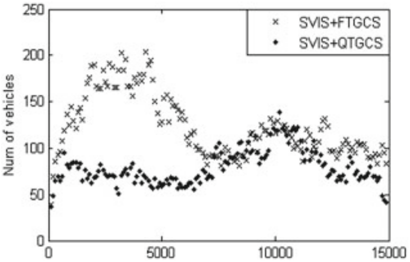


Fig. 7. The quantity of vehicles in traffic system

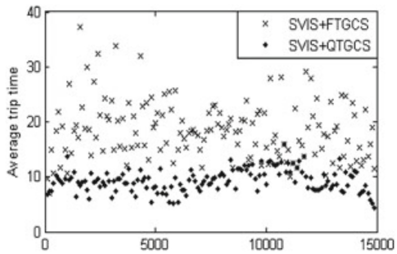


Fig. 8. Intersection vehicles average travel time in the lane

From Figs. 7 and 8, we can observe that the evaluation data obtained in SVIS and QTGCS collaboration is better than SVIS and FTGCS. The experimental results show that the FTGCS can improve the efficiency of traffic system and reduce vehicle travel time, compared with the traditional fixed-time strategy, based on the QTGCS can using real-time information of the road network, deploy green light of traffic light reasonably, and shorten vehicle travel time and delay time.

In SVIS, vehicle path selection based on induction information, and through the fuzzy logic control according to vehicles information to optimize timing scheme (FQTGCS), the light timing based on fuzzy optimize action. The evaluation data of SVIS and FQTGCS were compared with the evaluation data obtained by SVIS and QTGCS. The quantity of vehicles in traffic system is shown in Fig. 9, and vehicle average travel time in the lane is shown in Fig. 10:

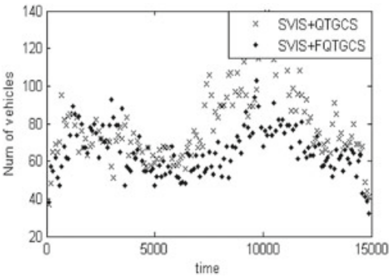


Fig. 9. The quantity of vehicles in traffic system

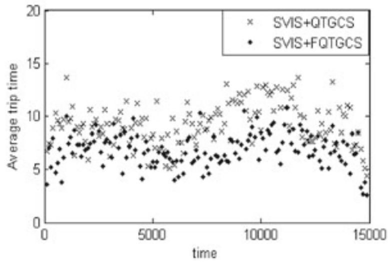


Fig. 10. Vehicle average travel time in the lane

From Figs. 9 and 10, we can observe that the evaluation data obtained the FQTGCS performance better than QTGCS. The experimental results show that the FQTGCS can improve the efficiency of traffic system and reduce vehicle travel time, compared with QTGCS. And the FQTGCS can using real-time information of the road network, deploy green light of traffic light reasonably, and shorten vehicle travel time and delay time.

4 Conclusion

In this paper, we formulated the traffic light timing strategy as a MDP problem, and applied Q-learning algorithm with Boltzmann based exploration strategy and fuzzy logic control strategy. A light learning agent will represent an intersection, where an agent can control the traffic light as much as road connected to the junction. The state will represent the congestion situation around the junction, it is seen from the traffic density of each phase in the intersection. And action represent which phase will get the green time to the light. Our algorithms are adaptive in nature and do not require road network model information, it can control traffic light in real-time by traffic information. Simulation results show that the performance of traffic system has been improved by the proposed traffic light timing strategy. In the current work, we only control traffic light with the information from joint junction without the light around. Thus, in the future, one could develop to coordinate the lights connected in a small area, and we intend to do our assumption in the future.

References

1. Abdulhai B, Pringle R, Karakoulas GJ (2003) Reinforcement learning for true adaptive traffic signal control. J Transp Eng 129(3):278–285
2. Arel I, Liu C et al (2010) Reinforcement learning-based multi-agent system for network traffic signal control. IET Intell Transp Syst 4(2):128–135
3. Azimirad E, Pariz N, Sistani MBN (2010) A novel fuzzy model and control of single intersection at urban traffic network. IEEE Syst J 4(1):107–111
4. Bi Y, Srinivasan D et al (2014) Type-2 fuzzy multi-intersection traffic signal control with differential evolution optimization. Expert Syst Appl 41(16):7338–7349
5. Chin YK, Wei YK et al. (2012) Q-learning traffic signal optimization within multiple intersections traffic network, pp 343–348
6. Chu T, Wang J, Cao J (2015) Kernel-based reinforcement learning for traffic signal control with adaptive feature selection, pp 1277–1282
7. Khamis MA, Gomaa W (2014) Adaptive multi-objective reinforcement learning with hybrid exploration for traffic signal control based on cooperative multi-agent framework. Eng Appl Artif Intell 29(3):134–151
8. Lu W, Zhang Y, Xie Y (2011) A multi-agent adaptive traffic signal control system using swarm intelligence and neuro-fuzzy reinforcement learning. In: Integrated and sustainable transportation system, pp 233–238
9. Moghaddam MJ, Hosseini M, Safabakhsh R (2015) Traffic light control based on fuzzy q-leaming. In: Artificial intelligence and signal processing, pp 124–128

Research on the Collaborative Governance of Innovation Network Based on the Extended JM Model

Chiyan Zou[1(✉)], Changyi Zhao[2,3], Tao Wang[1,4], and Xin Gu[1]

[1] Business School of Sichuan University, Chengdu 610064, People's Republic of China
inzcysist@sina.com
[2] Soft Science Institute of Sichuan University, Chengdu 610064, People's Republic of China
[3] Science Institute of Sichuan University, Chengdu 610064, People's Republic of China
[4] Institute of Innovation and Entrepreneurship Research, Chengdu 610064, People's Republic of China

Abstract. Orderly and efficient operation of the network security innovation, and actively promote the coordination effect of innovation network, is the key to further promote the economic development, is also an important way to achieve an innovative country. Based on the connotation of the collaborative governance, run, and the network structure of governance logic thinking and further analysis of the preliminary exploration to construct the basic framework of innovation network collaborative governance, in order to solve the existing problems such as insufficient innovation network effects play.

Keywords: Collaborative innovation · Innovation network · Inter-subjective relationship · Collaborative governance

1 Introduction

With the development of the economy, the model of single enterprise innovation gradually evolved into cooperative innovation, further promoting the development of innovation network. Collaborative innovation is the key impetus to promote the development of economy, it is complicated innovate organization. Cooperating the innovation network is a process that main bodies take the flowing of knowledge as the carrier from communication, coordination and cooperation to collaborative [15]. Therefore, the external formal of collaborative innovation is the development of innovation network, to realize the collaborative innovation is the ultimate purpose to form the innovation network.

Through the various innovative elements (micro and macro elements) innovation network use different collaborative ways to achieve nonlinear interaction

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_148

and coupling, coming into being effective knowledge flowing and optimized allocation in the self-organizing mechanism, the coupling mechanism, the network drive mechanism and coordination mechanism, to produce the overall effect is greater than the sum of the parts. In innovation network resources knowledge and behavior characteristic of subjects have great difference. There often appears the phenomenon such as fuzzy relations and negative interaction in the operation of the network. Perrow [8] proposed in a tight and complex system, the complex and unpredictable interaction between non-standard parts (namely different standards or specifications or character of the individual) will bring disastrous threat. Some integrate both parts of the performance equation-process and productivity from collaboration [6,7]. Throughout the domestic and international related researches, these mostly focus on collaborative innovation performance, operating carrier (mainly is knowledge), running mechanism and the role of subjects. Although some scholars proposed to attaches great importance to the network of relationships and interactions between subjects, but not for an in-depth studying, and about the process of innovation network operation controlling and synergies playing is not enough. In order to achieve the goals of innovation network, orderly and efficient operation of network organizations is necessary. How complicated and challenging collaboration can be, even though it may be needed now more than ever [2]. So we need to do more. Collaborative governance can effectively guarantee the network relationship, benign interaction and effective synergy [21], more suitable for the process controlling of innovation network operation.

The general governance focuses on network governance in a single point, line or dimension, ignoring the design and planning of overall network governance. However, collaborative governance emphasizes the integration of the network structure, the relationship and interaction [29]. Finally, "synergy" includes multiple subjects participating and cooperating, and integrity, interactive and sustainability; "Governance" is aimed at the network nodes, the relationship of network and the overall network, also includes the control of the governance target and the mechanism. Therefore, collaborative governance can guarantee network efficient operating. In view of this, this research refers to the thought of collaborative governance and try to build collaborative governance framework in innovation network, to better solve the problems in the innovation network operation, providing a reference for the management and development of innovation network.

2 The Connotation of the Collaborative Governance and Implementation

2.1 The Connotation of the Collaborative Governance

Collaborative governance theory comes from the coupling of synergy theory and governance theory. The collaborative governance study has got much more attention and is still in hot with different points, including highlight government administration [13,28], the multi-agent participating, and the multi-level cooperation governance [17,22]. Recently, studying on politics and environment performance by collaborative governance has blossomed, which concluded that different corporation tensions came different effective [1,3,26]. The UN commission on global governance in 1995 stressed that collaborative governance is a continuous process of reconciling different conflicting stakeholders and taking united action [11]. Jin [16] analyzed the collaborative governance of sublimation, and pointed out that in the era of the Web2.0 collaborative governance focus on system innovation, emphasizing the public power in the participation in the reasonable allocation between subjects, to break the shackles of the traditional single center model, realize the development in governance. On the basis of previous studies, this paper argues that "collaborative governance" is the multiple stakeholders in the governance mechanism establish cooperation, and through the network members multilateral interacting and coordinating to solve conflicts, control of conflicts and build cooperation in the advantageous network environment, finally, to guarantee the orderly operation and realize the goal of innovation network.

2.2 Implementation of Collaborative Governance

A 'Collaborative governance' model should include mechanisms for contributing more systematically, consistently and transparently [23]. About mechanism, studies include trusting and sharing mechanism, interactive process controlling and innovation property allocation mechanism [20,27]. Zhang [33] put forward that the collaborative model emphasizes to the idea of collaboration as the dominant factor, advocated by the regulation and government guidance to construct multi-center collaborative governance system; stressed the microeconomic foundation building, and used cognitive guidance mechanism and the balance of interests distribution mechanism to fully implement the bodies' self-management and from two aspects of process and result to ensure synergy mechanism long-term operating. Deng et al. [14] considering with the whole network, discussed collaborative governance mechanisms and analyzed the interaction between these mechanisms. Han et al. [32] concluded that the governance of innovation network not only need to pay attention to the subjects' relationship, but also to the effective guidance and governance of the communication and interaction between subjects, to provide scientific and efficient operation, policy, economic and technical environment for enterprise collaborative innovation network. Based with the previous researches, this paper holds that the implementation of the collaborative governance supported and safeguarded by collaborative governance mechanism, through guiding and controlling the relationship between network nodes, the network interaction, coordination effect and the interaction of the three tries to achieve self-management and overall coordination effect optimization.

3 The Logical Framework of Innovation Network Collaborative Governance Based on Expending JM Model

The relationships, interaction and collaboration is namely the logic of collaborative governance. Johanson and Mattsson [18] put forward to the relationship and interaction model (JM model), it organically linked the coupling relationship between network nodes and their interaction behavior. Sun [24] extended JM model on this base. He thought the relationship and interaction also involves the logic of network governance, and further to expand JM model from the two-dimensional structure to three-dimensional structure which also included interactive results-the collaborative (governance performance). Extended JM model provides a new perspective for innovation network governance. For this, we need to define the connotation of relationships and interactions in innovation network, as well as the possible synergies.

3.1 The Relationships Between Network Node

Granovetter [12] stressed that in the market, there are Strong relationship (Strong Ties), Weak relationship (weak Ties) and Free trade (Arms-length Ties) three forms among subjects, the Strong relationship means the subjects are more passionate and interactive, have high-integrity, frequent-interaction resources exchange. However, weak relationship is on the contrary. The degree of embedding in actual innovation network (strength of relationship between different subjects) and the network density (the ratio between network nodes connection number and maximum connection number) can influence the flow of knowledge in network, thus influence on the effect of collaborative innovation [30]. Michelfelder and Kratzer [19] analyzed that strong relationship with embedded characteristics, deepen relations and strengthen interaction mainly through the necessary commitment and contract. On the one hand, commitment encourage special assets investment to curb opportunism behavior; through mutual understanding, on the other hand, as well as the basis of continuous interaction, to promote trust based on emotional and relationship regulation emergence, to enhance internal convergence. Fliervoet et al. [10] analyzed collaborative governance through social network and proved that the relatively high network density improves the potential for collective action and collaboration. Strong relationship emerging based on weak relationship, weak relationship make network maintain high independence and flexibility, thus easier to expand relations connecting.

Innovation network continuously evolve in such main aspects as knowledge creation, diffusion and use [9]. Essentially, it is a kind of social relation network. Network relying on knowledge demands to establish relationship is weak relationship, which mainly is interests-oriented to make the network operation. With the interaction is increasingly deep, the cooperation presents diversification, the role of subjects shows the dynamic and diverse. Finally, weak relationship evolves to strong relationship to achieve greater knowledge sharing. In terms of knowledge's acquisition and integration, strong relationship promotes tacit knowledge flowing, and weak relationship promotes explicit knowledge flowing. Low cost characteristics of weak ties can absorb new knowledge and resources more, but generally causing the low-stable dominant network, insufficient knowledge flowing. Strong ties guarantee stable network, the cooperation of high willingness and sufficient knowledge flowing, strong relationship. Of course, too much strong ties will restrict the development of the network and the spread of knowledge. For these, we need to have different priorities. For most of China's national high-tech industry base, the high density of innovation network, for example, should be properly promote weak relationship evolve to the strong relations, designing and building complete industry chain and technology chain, to achieve higher conversion rate and win-win among the members. Meanwhile, it should be paid attention to the weak relationship developing to obtain more knowledge, especially the heterogeneous knowledge. In this way, in innovation network weak relationship strengthen the heterogeneous knowledge acquisition, strong relationships improve the efficiency of knowledge sharing.

3.2 The Interaction in Network

Interaction is another important dimension to collaborate in innovation network. It is mainly reflected in knowledge sharing and learning, the rational resources' allocation and its optimization, and so on [4]. Knowledge sharing is one of the process of knowledge flows, knowledge flow efficiency directly affects the synergy of innovation network [25]. In innovation network, for the difference of knowledge quality and quantity among subjects, and the difference of division, the complementary of knowledge, there is the relationship between knowledge supply and demand. Interactions between the main body, causing continuous circulating knowledge in the network between different subjects, to accelerate the process of collaborative innovation. Knowledge flows is one of the most common form in innovation activities, also is important activities to bring value-added. Therefore, this article regards knowledge flows as interaction in innovation network. In addition to the knowledge sharing, the knowledge flows also include knowledge transfer and knowledge creation. At the early stage of the innovation network operation, knowledge sharing is given priority, namely heterogeneity knowledge exchange among the subjects, so as to realize sharing. Knowledge creation stage is important process of creating new knowledge. Knowledge flows create knowledge synergy, and bring greater knowledge advantage for the whole network.

Competition and cooperation are the two basic forms of interaction in social [31]. In network, the cooperative interaction produces synergy effect, encourages members to pursue common interests, and promotes the knowledge spread effectively in network; benign competition interaction creates a new combination of the knowledge and value, through mutual learning and using knowledge, to reduce opportunism; Malignant competition interaction, however, causes contradictions and conflicts. Finally, hinder the relations developing, further hinder the synergistic effect. The two former increase the degree of familiarity and understanding between subject, reduce the communication cost, and avoid information asymmetry, low knowledge matching degree and the bridging degree for unfamiliar subjects, especially in knowledge.

Interaction of collaborative governance logic stresses the main body of governance is not a single subject, but all the members in network. They work together to promote the development of the network. Based on stakeholder theory [5], multiple participation not only promotes interaction, but also control and guide malignant it. In this way, it deepens understanding and trust among subjects and improves the efficiency of knowledge and resources flowing. Such as regional economic competition forces the Yangtze river delta to seek deep interactions with surrounding areas and other economic belt, and according to the geographical relations and the level of economic set up different levels of city circles. At the same time, improving city circle cooperation interaction, and promoting the development of regional economy.

3.3 The Result of Relationship and Interaction-Collaboration

In the process of interaction, new weak relationship emerges, further to evaluate strong relationship. Similarly, the benign interaction generates positive synergy, and further to strengthen the relationship. The strong and the weak relationship between network nodes, the benign interaction and positive synergy is in favor of achieving the goal together. To be sure, because of the differences of resources and capabilities between subjects directly leads to the result of interactive complex and uncertain and the evolution of and relationship, the collaboration created by relationship and interaction is diversity, dynamism and directivity. The collective rationality of repeated game formed the positive collaboration improved by Pareto, which is the balance of the relationship and benign interaction can generate "social facilitation effect", and add knowledge value, optimize innovation network environment and realize the goal of innovation network. On the contrary, due to the unbalance relationship and vicious interaction, there comes negative synergy, and leads to "social-loafing". Those cause the failure of innovation network targets with the subjects' negative behavior, deceptive interaction, vices and low negative effects.

In innovation network, the demand of knowledge causes to produce relationship, relationship promotes interaction, interaction promotes synergy, in

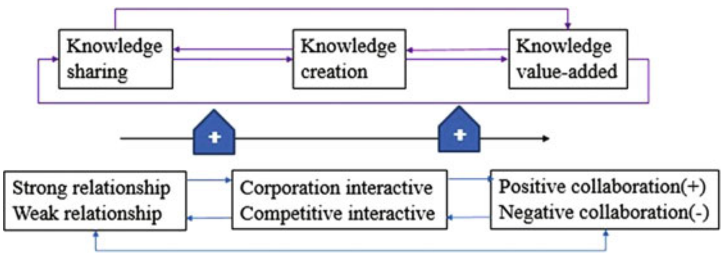


Fig. 1. The collaboration governance logical framework based on expending JM model

turn, synergy impels interaction and make organizational relationships deep. The whole process accelerates knowledge flow, promotes organizational power. Further, Organization network is strengthened. Those promote more subjects to participate in innovation, finally expanding the source of knowledge and the synergy beneficiary, also improving the efficiency of innovation network. New knowledge demands further to build new relationships and lead to a deeper interaction. It is a constant, circulating and spiral-rising process, as shown in Fig. 1, constituting a logic framework of innovation network collaborative governance.

4 Safeguard Mechanism of Collaborative Governance

The another key to success collaborative governance innovation network, is collaborative governance mechanism can forward drive relationship, interaction and coordination, ensure synchronous interaction and the orderly and efficient collaboration. as shown in Fig. 2, is the basic framework of innovation network collaborative governance.

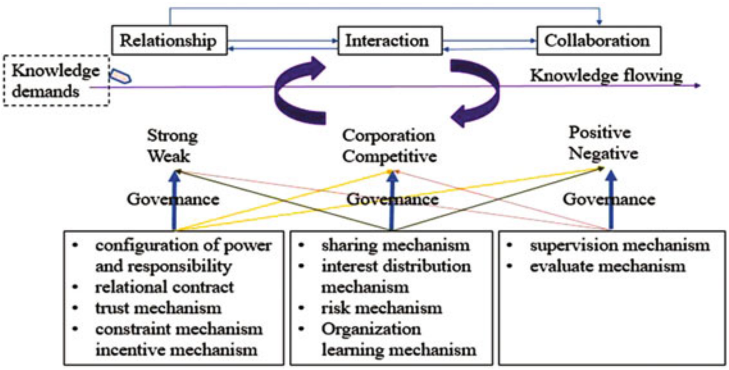


Fig. 2. The basic framework of innovation network collaborative governance

First, the configuration of power and responsibility, relational contract, the trust mechanism, constraint mechanism and incentive mechanism promote weak ties form, further to the strong relationship evolution, meanwhile, to control these two relationships, to make up the malpractice and the insufficiency, to take their advantages, to ensure maximum effect. Accrual configuration make the government, enterprises, universities, research institutes and intermediary service organizations these subjects fully play function advantages, fully coordinate. At the same time, overcome the limitations of self-regulation and government single coordination. Relational contract and trust mechanism prompt relationship formation and strengthen, so as to guarantee the stability of the network, to promote knowledge and resources fully flow and exploit. Among them, the relational contract can product weak relations, to maintain the network openness and the heterogeneity of knowledge, to develop the network value. Trust must be embedded network, could it exist. To product and strengthen the relationship can establish and strengthen trust; On the contrary, the trust is the necessary condition to build strong relationships. On the one hand, Constraint mechanism and incentive mechanism is in favor of reducing the cost of strong relationship, and the contradiction caused by complicated relationship, On the other hand, to encourage the network extension, improve network activity.

Second, establishing sharing mechanism for knowledge and other resources, interest distribution mechanism and risk mechanism to avoid opportunism. Being satisfied subjects with interests' needs, and promoting cooperation and benign interaction, promoting the knowledge's efficient flowing in the network. As well as principal partners in collaborative innovation network, subjects also is competitors, benefit disputes and conflicts among them will affect the overall effect of the collaborative governance. Meanwhile, their some characteristics can affect the formation and development of innovation network collaborative governance. As a result, interest distribution mechanism and risk mechanism reasonable guarantee interest allocation effective and fair. The conclusion of the agricultural model, when knowledge's gap is very big, strong party interaction's will is insufficient, combined with the weak absorption ability is limited, leading to a lower efficiency of interaction. Conversely, when the gap is small, the effective is better. Sharing mechanism which can reduce the knowledge's gap, and can avoid the information asymmetry. In the process of knowledge flows, involving the subjects learn and internalized knowledge. Therefore, organizational learning mechanism needs to further strengthen, to reduce the knowledge differences between subjects, so as to promote knowledge flow.

Third, through establishing supervision mechanism and evaluate mechanism, evaluating subjects' execution and the completion of target. Due to the diversification of innovation network, supervision mechanism is hard to effective operate. For effective supervising the whole process of collaborative governance, firstly, Institutionalized problems such as the organization's powers and the constitution and the profit distribution, and implement the innovation network. Increasing default cost and effectively reducing the coordination costs, to constraints subjects' behavior with institutional in collaborative governance. Second, strengthen moral construction, forming public opinion and supervision mechanism. Supervising subjects' behavior in the process of collaborative governance with moral. Through the evaluate mechanism to clear synergy governance effect, clear and control each problem source in every stage and every node.

Of course, the cooperative governance mechanism is not only for one link in the relationship, interaction and synergy, but on the whole process governance. Constraint mechanism, for example, can also constraint subjects' behavior in the collaborative governance, control interaction. Supervision mechanism and evaluating mechanism is also valid for each link and each movement. And the characteristic of innovation network and collaborative innovation represents the different phase characteristics, cooperative governance mechanism should also be adjusted. In the process of the collaborative governance, collaborative governance logic is the basis of collaborative governance, also is the junction between innovation network and collaborative governance. Collaboration governance mechanism is support and assurance of the collaborative governance. All of them mutually dependent and interacting, forming the prerequisite and guarantee of innovation network collaborative governance, driving the network orderly and efficiently operating, finally to achieve maximum synergy innovation effect.

5 Conclusion

This paper mainly discusses how to solve the problems in the operation of innovation network through cooperative governance, according to the extension of JM model further analysis of the collaborative governance logic, on this basis to analyze the innovation network of cooperative governance, and put forward the safeguard measures of innovation network collaborative governance, so as to support and guarantee for the operation of the collaborative governance, to promote innovation network synergy. Effective cooperative governance mechanism is able to efficiently governance the relationship, interaction and synergy in innovation network. Finally, to solve the relationship which is not clear and soured between the main body in innovation network, network is running off track, collaborative innovation is unsustainable, coordination is inefficient and so on. In the future, further research on collaborative governance mechanism and the strength of the control network relations, and the definition and controlling for the relationship intensity in governance model etc.

Acknowledgements. This paper is supported by the National Natural Science Foundation of China the Research on Knowledge Chain Governance Based on the evolution of inter organizational relationships: (No. 005020540109), and Study on informal governance of enterprise innovation network (No. 71602132), and the Central University Foundation of Sichuan University: the Research on the informal governance of Chinese enterprises' collaborative innovation network under the new normal condition (No. kzx2015-sb49).

References

1. Ansell C, Torfing J (2015) How does collaborative governance scale? Policy Polit 43(3):315–329
2. Bryson JM, Crosby BC, Stone MM (2015) Designing and implementing cross-sector collaborations: needed and challenging. Public Adm Rev 75(5):647–663
3. Charron N, Dijkstra L, Lapuente V (2014) Regional governance matters: quality of government within european union member states. Reg Stud 48(1):68–90
4. Chen J, Yang YJ (2012) Theoretical basis and content for collaborative innovation. Stud Sci Sci 62:270–277
5. Dugger WM (1996) The mechanisms of governance. J Econ Issues 44(4):261–281
6. Emerson K, Nabatchi T (2015) Collaborative governance regimes
7. Emerson K, Nabatchi T, Balogh S (2012) An integrative framework for collaborative governance. J Public Adm Res Theor 22(1):1–29

8. Essay (1979) Complex organizations: a critical essay. McGraw-Hill companies, Incorporated
9. Etzkowitz H, Leydesdorff LA (2001) Universities and the global knowledge economy: a triple helix of university-industry-government relations. Pinter
10. Fliervoet JM, Geerling GW et al (2016) Analyzing collaborative governance through social network analysis: a case study of river management along the waal river in the netherlands. Environ Manage 57(2):1–13
11. Commission on Global Governance (1995) Our global neighbourhood: the report of the commission on global governance. Oxford University Press, Oxford
12. Granovetter M (2005) The impact of social structure on economic outcomes. J Econ Perspect 19(1):33–50
13. Halachmi A (2005) Governance and risk management: challenges and public productivity. Int J Public Sect Manage 18(4):300–317
14. Hong Q, Zhang H, Business SO (2016) The structure characteristic and self-organizing evolution of collaborative knowledge innovation network. J Huaiyin Inst Technol 317:41–49
15. Jin C (2011) Collaborative innovation and construction of national scientific research ability. Sci Manage Res 12:1762–1763
16. Jin T (2016) Collaborative governance innovation web2.0 era: the system of governance. Academics 324:5–14
17. Agranoff R, Mcguire M, Joaquin E (2005) Collaborative public management: new strategies for local governments. J Federalism 35(4):637–639
18. Johanson J, Mattsson LG (1987) Interorganizational relations in industrial systems: a network approach compared with the transaction cost approach. Int Stud Manage Organ 17(1):34–48
19. Michelfelder I, Kratzer J (2013) Why and how combining strong and weak ties within a single interorganizational R&D collaboration outperforms other collaboration structures. J Prod Innov Manage 30(6):1159–1177
20. Panico C (2011) On the contractual governance of research collaborations: allocating control and intellectual property rights in the shadow of potential termination. Res Policy 40(10):1403–1411
21. Peng X (2010) Three-dimensional governance, relationship management, network management and knowledge management, knowledge partner relationships between network organization governance research. J Libr Intell 6:121–126
22. Powell WW (1996) Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology. Adm Sci Q 41(1):116–145
23. Prideaux M (2016) Considering collaborative governance: Nongovernmental organisations, the convention on migratory species and the convention on wetlands of international importance. Wyoming geological association
24. Sun G (2013) The relationship, interaction and collaboration, the logical of network-organization. China Ind Econ 11:14–20
25. Tu ZZ, Gu X (2013) Study on process of industry-university-research institute collaborative innovation based on knowledge flow. Stud Sci Sci 39:182–193
26. Ulibarri N (2015) Tracing process to performance of collaborative governance: a comparative case study of federal hydropower licensing. Policy Stud J 43(2):283–308
27. Uzzi B (1997) Errata: social structure and competition in interfirm networks: the paradox of embeddedness. Adm Sci Q 42(2):35–67
28. Warren ME (2011) Foundations and frontiers of deliberative governance. Oxford University Press, Oxford

29. Wu C, Zhuang Y (2013) Cooperative governance, the key variables, influencing factors and implementation ways. J Theor Explor 3:73–77
30. Yubing HE, Zhang Y (2015) Research on coupling of network types and the collaboration modes of industry-university-research institution collaborative innovation. Sci Sci Manage S&T 9:79–85
31. Zeng Z, Tu Z (1993) The heart of the collision, ethics-sociology of the virtual and real. Hunan publishing, Changsha
32. Zhang S, Yuanjia HU et al (2011) Collaborative innovation model of traditional Chinese medicinal pharmaceutical enterprises. Sci Technol Manage Res 27:42–58
33. Zhang Z (2015) On the generation logic and the construction path of collaborative governance. Chinese public administration

Performance Evaluation of Housing Price Regulation Policy in China: Based on ARIMA Model and Intervention Analysis

Chang Liu[1], Yixiao Zhou[2], Wei Zhao[3], Qiang Jiang[3(✉)], Xuedong Liang[3], Hao Li[3], Hua Huang[3], and Shucen Fan[4]

[1] School of Securities and Futures, Southwestern University of Finance and Economics, Chengdu 610052, People's Republic of China
[2] School of Economics and Finance, Deakin Univesity, Perth, Australia
[3] Business School, Sichuan University, Chengdu 610065, People's Republic of China
jiang.qiang@outlook.com
[4] Department of Economics and Management, Sichuan Technology and Business University, Chengdu 611745, People's Republic of China

Abstract. Chinese housing market is rapidly growing up. However, the study on housing market regulation is inadequate. In this paper, housing price indexes from 2006 to the present are analyzed based on the ARIMA model and intervention analysis, and the change of this index can reflect the effective date and valid date of country's housing price regulation policy. This paper established systematically the quantitative analysis model, studied the timing effects of new regulations, and improved the quality of the house price index by using the monthly data. Based on the model results, we evaluated the performance of the regulation policy of housing price on three criteria including anticipation, conductivity and additivity. This can provide the field evidence for policy makers to design better housing policy.

Keywords: ARIMA model · Intervention analysis · Quantitative analysis · Housing price regulation policy · Performance evaluation

1 Introduction

Since the 30 years of reform and opening up, China's real estate industry gradually changed from the planned economy system to market economic system, and real estate and related industries have been the most important industry in China's regional social economy and played an important role in the national macroeconomic development. However, along with the gradual progress of the marketization process of the housing market, we are facing some unexpected

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_149

negative effects of China's housing market including high housing prices and low housing affordability, housing prices increased too fast, the fever of land speculation. These negative effects demonstrated that the housing market operation mechanism has some limits. The direct consequence is that houses have become unaffordable for the low-income groups being not satisfied and housing structure becoming unbalanced.

To solve the current problems in China's housing market, the government introduced a series of monetary policy and fiscal policy. The policy makers tried to change the performances of banks and real estate developers in the market and cool the overheated housing market by standardizing housing market laws and regulations. Especially after 2008, the rapidly rising housing price has become the focus of public opinion, how to keep housing price away from rapid increases has become a vital problem which the government needs to solve urgently. Under this background, a series of price regulation policy were introduced at the historic moment.

Since 2005, housing price regulation policy which issued by Chinese government can be roughly divided into two categories: First, policy instruments are design to suppress housing prices, such as issuing purchase orders, increasing bank interest rates, raising housing taxes, raising the down-payment ratio of real estate and so on. Second, policy instruments are design to stimulate the housing prices, such as: tax reductions, lower down-payments, lower interest rates, and so on.

Since 2005, little evidence has been found in the efficiency of the existing housing policy. As the housing market policy failure in 2005, China's housing market continued to deteriorate and housing prices in many cities increased sharply in 2006. From 2005 to 2007, the national housing policy focused on increasing bank interest rates and raising tax rates. For example, in May 2006, the State Administration of Taxation issued a notice on strengthening housing business tax levy management. This new tax regulation required that after June 1st, individuals housing purchased less than 5 years shall paid the business taxes at the full rate; In 2007, the People's bank of China increased interest rates by 6 times and the deposit reserve rate by 10 times, it is still failed to control rapid increased property prices.

In 2008, to cope with the financial crisis, the Chinese government decided to stimulate China's housing market rather than suppress it. Therefore, the government launched a series of new policy to stimulate housing prices. These new policies included the temporary exemption of stamp duty on personal sales or purchase housing; the temporary exemption of land value-added tax on personal sales housing; local government can introduce their own policies to encourage housing consumption tax relief and so on. In 2009, the housing price regulation policy was mainly on the inheritance, strengthening and deepening of 2008 real estate regulation policy, while highlighting the focus of strengthening the housing security. Therefore, the early 2009 housing price regulation policy mainly includes several aspects like tax, financial market and housing security. In 2009, while most of the industry was facing a financial crisis, China's real estate industry had achieved a substantial development. However, these stimulation policies since 2008 left a hidden danger to the subsequent regulation. After 2010, the government began to realize that the real estate market may be out of control, so it was very difficult to curb its rising momentum again. In theory, if a series of policies issued by the government can be effective, China's housing prices may change to the expected direction as the policy makers' design. However, policy effects are not obvious in the China's housing market performance.

In this paper, by using the ARIMA model and intervention analysis, we analyze the effectiveness of price regulation in China. So researchers and policy makers can understand the critical issues in China's housing policies including the effectiveness, failure, superposition and neutralization of housing policies. In general, this study can provide policy recommendations for the design of a more effective housing price regulation.

This paper has been organized in the following way. Section 2 begins by laying out the theoretical dimensions of the ARIMA models and previous research in this field, and looks at how ARIMA models can be applied in the China's housing policy research. In Sect. 3, we conduct the empirical model by exploring the ARIMA model and establishing the intervention analysis model to see the results of the housing price regulation. In Sect. 4, we applied the ARIMA model and intervention analysis into China' housing market. Section 5 presents the findings of the research, focusing on the three key themes including the effectiveness of policy, the validity of policy, the neutralization and superposition of policy. In Sect. 6, we present the conclusions about the effectiveness of policies, the validity of housing policies and the superposition between the policies.

2 Literature Review

There are very rich economic research results about national regulation policy effect on housing prices. The existing studies mainly focus on four aspects:

First, it is the analysis of several factors affecting the housing prices to discuss the effect of regulation policies, Such as Wang and Chen [14] in Wuhan City as an example, the author thought that four indicators: sales accounted for below 90 square meters, the new supply area, the money supply (M2), the lending rate, put a significant impact on Wuhan housing prices. This study focuses on the qualitative analysis of the regulatory policy and its effect, but it is lack of accurate interpretation to the effect of regulation policy.

Second, it is the single regulation policy effects on the real estate prices, for example, monetary policy, land policy and tax policy. Nie and Liu [11] thought that China's monetary policy can affect real estate prices, and the money supply was more significant than the interest rate policy regulation function by analyzing the monetary policy from 2005 to 1994. At present, most of the domestic literatures focus on qualitative analysis to study the effect of the policy on the regulation of real estate market, and most research results show that monetary policy has a significant impact on housing prices. This study did not consider the regulation effects of a number of policies under the joint action comprehensively, and be lack of a comprehensive research on overall coordination and interaction.

Third, it is to use a social science or econometric model to analyze, such as Wu [17] based on PSR model to establish housing price regulation policy effectiveness evaluation system, after the solution of pressure, state and response index weight system, through the analysis of the relationship between the three indexes, found that housing price regulation policies in recent years in 2005–2006, and 2011–2012 were relatively effective, proposed when country introduced a more severe structural adjustment policy can make effective control on housing prices.

Fourth, researchers are interesting in evaluating the policy performance. Chen and Fang [3] explored VECM and DSGE model and found that policies on restricting house purchase are mainly directed at speculative demands so that they are suitable for house price control of first-tier cities, while the down payment ratio policy on mortgage is directed at rigid demands and they are fit for non-first-tier cities. Chen W [2] applied a multiple-factor panel data model to show that a high degree of market failure is associated with a high ratio of persistent components in the gap between price and equilibrium. Chen found the housing policies did improve the housing market's efficiency. Feng and Wu [8] studied whether there is an asset bubble in China's residential housing market by using a an equilibrium asset-pricing approach. They did not found the existence of a house price bubble at the national level.

Fifth, non-economic factors such as education showed more impacts on housing prices. Agarwal et al. [1] studied the relation between school allocation rules and housing prices in Singapore by using a quasi-experiment. They found that private housing prices within 1-km zone and in 1-km to 2-km zone from the old school zone decline by 2.9% and 6.0%, respectively, 6 months before the school relocation events. Feng and Lu [7] studied if school quality will affect housing prices in Sanghai, China. They found that a good school may increase housing prices by 17.1%. Wen et al. [15] reported the similar results by using the data from Hangzhou, China. The found that elementary and junior high schools have a significant effect on local housing prices. Furthermore, they reported that the housing price increases by 2.737% or 0.904% when the house is located less than 1 Km from the high school or college.

ARIMA has been an effective approach in the housing market study. Woo and Sung-Suk [16] found ARIMA is a useful in exploring the relation between housing prices and interest rates. Previous studies used logistic model [5,10], GARCH model [4,21], canonical correlation analysis and other methods to analyze the regulation effect of comprehensive policy, mainly through analyzing the change of housing prices, land prices and other major indicators to judge the effectiveness of regulation policy and evaluate the regulation effect of comprehensive policy. Although the research analyzed the overall effect of the real estate regulation policy, there was no analysis of the mutual influence between different time point and type policies.

In recent years, there is a growing research interest from the management science field. In the agent theory, Zhang et al. [19] studied Beijing's housing market by using a multi-agent model. Ge [9] applied another agent model in American housing market. Neural Networks are another popular approach. Lim et al. [13] applied Neural Networks in housing prices predications. Other studies included index theories [6,12,18,20].

In summary, the current study mainly focused on the qualitative study of single and comprehensive regulation policy effect, and there is less literature to use quantitative analysis method to study the regulation policies the superposition and neutralizing effect between regulation policies. Therefore, this article adopts the monthly price indexes between January 2006 and February 2015 to establish housing price regulation policy effectiveness evaluation system based on the ARIMA model and intervention analysis model. Furthermore, this study analyzes the influence of policy factors on the real estate prices empirically. According to the results of the model, we conducted a quantitative performance evaluation on house price regulation policy by years. The quantitative study focuses on three key themes including analyzing the anticipation, conductivity of housing price regulation policy implementation effect in recent years, as well as the mutual influence between different time point housing price regulation policy implementation effect, their hedge and superposition, and reasonable predictions of the effects of policies. These results can provide a scientific basis for the formulation of the new regulation policies.

3 Theoretical Basis

3.1 ARIMA Model

The basic idea of the ARIMA model is to regard the data sequence as a random sequence which is formed by the time lapse of the forecast object, which uses a certain mathematical model to describe the sequence. Once recognized, the model can predict the future value from the past and present value of the time series. The ARlMA model in economic forecasting process not only takes into account the dependence of economic phenomenon in the time series, but also considers the interference of the stochastic volatility, which has a higher accurate rate to forecast the economic operation of short-term trend and is one of the widely used methods. Because the economic data is often autocorrelation and non-stationary time series, and the ARIMA model can effectively deal with autocorrelation and non-stationary data, while the ARIMA model is stable that can effectively combine with the intervention model, the ARIMA model is used to predict the housing price index of our country.

Between January 2006 and February 2015, the monthly housing price index can be seen as a random time series which is formed over time. By analyzing the randomness, stationary and seasonal factors of the housing price index of this time series, using ARIMA mathematical model to describe the correlation or dependence between the values of these single month housing price index. Thus taking advantage of the past and present housing price index to predict the future index, it can achieve the purpose of predicting the influence of housing price regulation policy on the future price.

3.2 Intervention Analysis

Time series are often influenced by special events and situations such as external interventions. Intervention refers to the stand or fall of forecasting model fitting degree, that is the pros or cons of fitting degree between the simulated value generated by the prediction model and the historical actual value.

(1) Intervention variables' form

The basic variables of intervention analysis model are intervening variables. There are two kinds of common interference variables: the first is the sustainable intervening variables, which indicates that there have been affected when T moment happened, then the step function representation can be used; the second is transient interference variables, which expresses that it can been affected just for the moment when it occurs at a certain moment, using unit impulse function. This article not only studied the regulation effect of housing policy at a certain time, but also analyzed the impact of these policies on the future housing prices, so we chose the first kind of intervention variables' form, as Eq. (1):

$$S_t^T = \begin{cases} 0, \text{before intervention events } (t < T) \\ 1, \text{after intervention events } (t < T). \end{cases} \tag{1}$$

(2) Intervention events' form

In the process of policy implementation, the intervention events may have an impact on the regulation effect. The influence of this kind of intervention events may suddenly start, and last for a long time. Set intervention's influence on the dependent variable being fixed, starting from a time T, but the extent of impact is unknown, that is, the size of the dependent variable is unknown. The intervention model of this effect can be written as in Eq. (2):

$$Z_t = \omega S_t^T \tag{2}$$

It is the unknown parameter of the intervention impact strength. While the rough stability series can be into stationary series by differencing, the intervention model can be adjusted in Eq. (3):

$$(1 - B)Z_t = \omega S_t^T, \tag{3}$$

where: B is a backward shift operator. If the intervention event has to lag a number of periods before impact, such as b periods, then the intervention model can be further adjusted in Eq. (4):

$$Z_t = \omega B^b S_t^T \tag{4}$$

where: $S_t^T = 1$,

$$Z_t(1 - \delta B) = \omega, \tag{5}$$

$$Z_t - Z_{t-1}\delta = \omega. \tag{6}$$

By deformation

$$Z_t = \frac{\omega}{1 - \delta B}, \tag{7}$$

where, B is a backward shift operator. According to the forecast results, if the fit of the forecasting model is higher, and the forecast value is also in line with the actual situation, then the prediction model has a certain application value. On the contrary, the prediction model is invalid. So on the basis of the ARIMA model, it is necessary to carry out the intervention analysis. The purpose of our intervention analysis is to assess the specific impact of the other policy intervention or emergency on the economic environment and economic process by using the quantitative analysis.

4 Empirical Study

4.1 Evaluation Index Selection

This paper investigates how the housing market response to the change of the housing price regulation. Thus housing price index is the most direct indicators to reflect the housing price changes. Housing price index is the advantage of "homogeneous", which can exclude the price fluctuations which due to market supply and demand relations and other reasons after the effect of these factors: housing quality, construction structure, geographical location and others. The monthly housing price indexes of 2006 to 2010 can be achieved in the iFinD software directly, while the 2011 to February 2015 index can be calculated by averaging monthly housing price indexes of 70 large and medium-sized cities nationwide. Thus the random time series of housing price index is obtained.

4.2 Evaluation Index Structure

(1) Establishing a single variable ARIMA model

The following figure is a China commodity house actual price index time curve graph (sampling frequency is once a month) from January 2006 to February 2015. It is sharply higher in 2007, 2010 and 2014, and housing price index fell to trough in 2008 (Fig. 1).

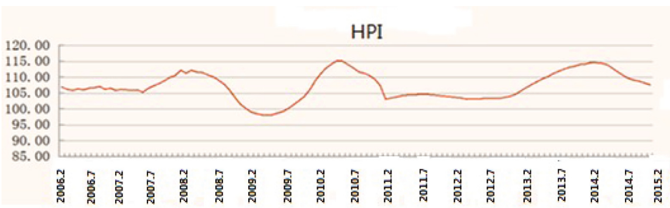


Fig. 1. Real estate price index

Observing the housing price index time curve in Fig. 1, we find that the data is not stable and contain a clear trend. Data is changed into a stationary random sequence by first order difference, as shown in Fig. 2:

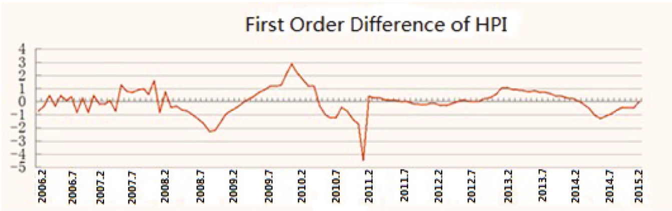


Fig. 2. Sales price index after first - order difference

After the normalization of the original data and repeated recognition of ARIMA model, we determine $p = 1$, $d = 1$, $q = 1$. Carrying out parameter estimation for ARIMA (1, 1, 1) model, we establish the univariate ARIMA model:

$X_t = 0.0190 + 0.8370X_{t-1} + \varepsilon_t 0.2787\varepsilon_{t-1}$, X_t is difference, ε is residuals.

Predicting the residual, the residual series time curve of the model is not a trend term, similar to random white noise, so it pass the model test (Fig. 3).

(2) Establishing intervention analysis model

We use the ARIMA (1, 1, 1) model to take extrapolating forecast for housing price index from January 2006 to February 2015, then use the actual value to minus the predicted value, the difference from which is the impact utility generated by macroeconomic regulation policy, remember to Z_t.

According to the form of selected intervention events, instability can be into a stationary series by differencing, and using housing price index from 2006 January to 2015 February, ARIMA predictive value and economic policy conflict utility, we can estimate the parameter of intervention model (Fig. 4).

By regression in Eq. (8),

$$Z_t = 0.8357Z_{t-1} - 0.0044, \tag{8}$$

where, Z_t is the difference between the actual values and predicted values. Namely in Eq. (9),

$$Z_t = \frac{0.0044}{1 - 0.8357B}. \tag{9}$$

Through the t-test of the regression value of δ, it is found that the estimated parameter δ is significant. So after the intervention analysis, the ARIMA model is in Eq. (10):

$$X_t = 0.0190 + 0.8370X_{t-1} + \varepsilon_t - 0.2787\varepsilon_{t-1} - \frac{0.0044}{1 - 0.8357B}. \tag{10}$$

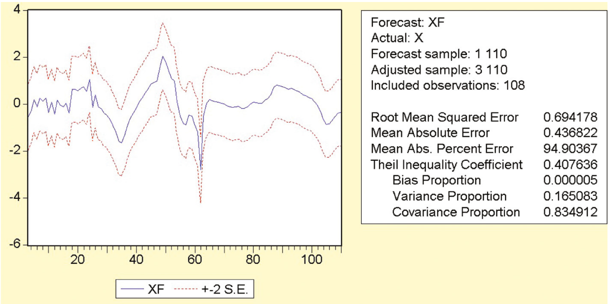


Fig. 3. Model testing

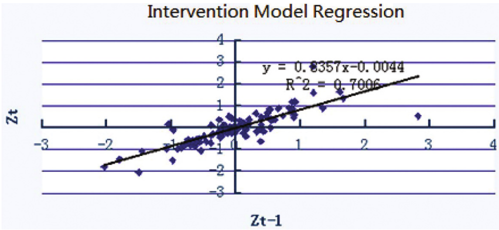


Fig. 4. Residual regression model

5 Results and Analysis

The study shows that the government's housing price regulation policy is not always effective, even if it is a utility policy, it may not take effect in time. We also found that the different policies will affect each other. The following paragraphs will analyze the effectiveness of policy, the validity of policy and the superposition between the policies.

5.1 Effectiveness of Policy

By analyzing the housing price regulation policy since 2006, we can find that there are three main kinds of policies which have a timely effect. First is involved in the adjustment of the relevant taxes. Such as in May 2006, the state promulgated the country of six, and the State Administration of Taxation issued the notice of the relevant issues on strengthening the management of housing business tax collection. This policy promulgation is St2. Taking it into validity test, we can see that the next month policy effectiveness is significant, which indicates that China's commercial housing price index has made the reflection in that month.

Second is to adjust the bank loan interest rate and the first loan ratio. In September 2006, the people's bank of China issued a notice on strengthening the management of the commercial real estate credit, which clearly stipulates that the loans down payment ratio of the second set housing shall not be less than 40%, lending rates may not be lower than 1.1 times of the benchmark interest rate which announced by people's Bank of China over same period and grade, and the proportion of the loans down payment and interest rates will be significantly improved with the increase of volume, the purpose of which is to regulate trade order and rectify the real estate market.

Third is about to investigate and punish the violations of law and discipline and the ownership disputes. However, the policy which can take effect immediately is still relatively less, and most of the policies are lagging, such as the state issued the regulations to strengthen land supply regulation and shorten the period of land development in October 2007, and in the same month, property tax "idling" ten spread to 10 cities. Housing price index began to make the corresponding reflection in December.

5.2 Validity of Policy

Among these housing price regulation policies since 2006, raising lending rates, introducing second-hand housing business tax policies, regulating loans down payment ratio on second homes and other relevant policies played a profound influence on housing price regulation, which has a long period of validity in a period of time after the introduction of the policies. Such as in April 2006.

Interest rates on deposits and loans were raised eight times by the people's bank of China, mortgage interest rates were up-regulated again regulation staged a comeback, the implementation of this policy had an indirect influence on the real estate prices, which continued to 2008; In September 2007, the central bank regulated that as a family unit, the loans down payment on second homes shall not be less than 40%, the interest rate shall not be less than 1.1 times of the benchmark interest rate, cracking down on speculation, which had played a certain influence on price regulation in later three years.

The property market regulation "new country of five" was introduced in the state council executive meeting, which reiterated its adherence to the implementation of the regulation policy which took the purchase limit and credit limit as the core, resolutely crack down on speculative investment, urged all localities to publish the annual housing price regulation targets. "New country of five" also stipulates that tax was levied on personal income of 20% in second-hand housing transaction. This provision would discourage investment demand and affects the volume, which is a positive effect for the real estate regulation, brought inhibitory effect on the residents' improved housing demand. The implementation of these policies had a long-term impact on the housing price regulation in China. And the introduction of the new policy in the recent two or three years, due to the implementation time is not longer, we can't see whether it has a long term.

5.3 Neutralization and Superposition of Policy

In the process of housing price regulation, in the face of different reflection of the market, the later policies released by government will cause a certain influence to the previous policies. This influence is two-sided, either playing a strengthening effect on the previous policies, call superposition effect; or having a conflict to previous policies, call neutralizing effect.

In July 2009, in order to strengthen the supervision of credit funds, banking regulatory bureau released “Interim Measures for the administration of fixed assets loans” and “guidelines for project financing”, to ensure that the fixed assets loan fund is truly used for the needs of the real economy, to prevent diversion of loans, and prevent the bank risk under the trend of rapid loan growth. In the next month of document issuance, the market did not make obvious reflection. We set this policy issuance as St23. In December, the State Council discussed and passed “opinions on the trial of the social insurance fund budget”. The views decided to make the period of the personal housing transfer business tax exemption from 2 years be back to five years. And five ministries and commissions such as the Ministry of finance, Ministry of land and resources, the central bank, the supervision published “notice on the further strengthening the management of land transfer revenue and expenditure”, improving the land developer down-payment. We se this policy as St25.

As you can see, after the promulgation of ST23, St25 began to change significantly, indicating that the market began to be sensitive to the policy enacted in July that year, that is, St25 had a superposition effect on St23.

However, in September 2010, with the introduction of the policy that the national commercial banks suspended issuing residents and families to buy the third and above housing loans (St27), the significance of the impact of the policy enacted in July 2009 (St23) on the market disappeared, which shows that the two policies have a neutralization. As shown in Fig. 5, the asterisk represents a significant effect.

	2009.7	2009.9	2009.12	2010.1	2010.9	2012.2	2012.3	2012.5	2012.6
St23				*	*		*		
St24					*				
St25					*	*	*		
St26									
St27							*	*	*
St28									
St29									
St30									
St31									

Fig. 5. Policy impact

6 Conclusion

This paper quantitatively analyzes the effectiveness of housing price regulation policy. Empirical analysis results show that mostly policies do not began to take effect in that month, and the policy regulation effect is much poorer in the second half of 2007, the first half of 2008, 2009 and 2012. The effective time of different policy is also different, where the policies have a certain impact on the next two years housing prices in September 2007 and February 2013. Policies launched by government may have overlapping or effect neutralization. These policies are international.

Using ARIMA model and intervention analysis to analyze the effect of the policy intervention, predicting the effect of this policy, the effect of superposition or neutralize is analyzed by linking the policy and previous policy. It will provide reference for the relevant functional departments to formulate effective housing price regulation, so that the policies have been introduced together to achieve the best regulation effect.

Acknowledgements. This paper is supported by the National Natural Science Foundation of China (General Program Grant No. 71473204), Sichuan Soft science research project (Grant No. 2015ZR0128), Chengdu Soft science research project (Grant No. 2015-RK00-00087-ZF), Chengdu Soft science research project (Grant No. 2015-RK00-00257-ZF), Research Grant Foundation of Sichuan Education Department (Grant No. xyzx1506) and Science Foundation of Ministry of Education of China (Grant No. 14YJC790053).

References

1. Agarwal S, Rengarajan S et al (2016) School allocation rules and housing prices: a quasi-experiment with school relocation events in singapore. Reg Sci Urban Econ 58:42–56
2. Chen WD, Hall S, Pauly P (2016) Policy failure or success? Detecting market failure in China's housing market. Econ Model 56:109–121
3. Chen X, Fang Y (2016) Research on the implementation and exit effect of the control policies on real estate based on VECM and DSGE model. Mod Econ Sci 38:31–43
4. Crawford GW, Fratantoni MC (2003) Assessing the forecasting performance of regime-switching, arima [auto regressive integrated moving average] and garch [generalized autoregressive conditional heteroskedasticity] models of house prices. Real Estate Econ 31:223–243
5. Dang Y, Liu Z, Zhang W (2014) Land-based interests and the spatial distribution of affordable housing development: the case of Beijing, China. Habitat Int 44:137–145
6. Dezhi L, Yanchao C et al (2016) Assessing the integrated sustainability of a public rental housing project from the perspective of complex eco-system. M.E. Sharpe
7. Feng H, Lu M (2010) School quality and housing prices: empirical evidence from a natural experiment in Shanghai, China. J Hous Econ 22(4):291–307
8. Feng Q, Wu GL (2015) Bubble or riddle? An asset-pricing approach evaluation on China's housing market. Econ Model 46:376–383
9. Ge J (2017) Endogenous rise and collapse of housing price: an agent-based model of the housing market. Comput Environ Urban Syst 62:182–198
10. Hui ECM, Zhong J, Yu K (2016) Land use, housing preferences and income poverty: in the context of a fast rising market. Land Use Policy 58:289–301
11. Nie X, Liu C (2005) An empirical study on the transmission of Chinese monetary policy toward investment. Manage Rev 17(2):51–54

12. Shi J, Zou W (2009) Forecasting house price index for USA with autoregressive integrated moving average models. In: The international conference on management of technology, Taiyuan, pp 590–594
13. Wan TL, Wang L et al (2016) Housing price prediction using neural networks. In: International conference on natural computation and fuzzy systems and knowledge discovery, pp 518–522
14. Wang L (2013) The effect of recent real estate regulation policy on house price: a case study of Wuhan city. Mod Prop Manage 221(2):306–315
15. Wen H, Zhang Y, Zhang L (2014) Do educational facilities affect housing price? An empirical study in Hangzhou, China. Habitat Int 42(42):155–163
16. Woo K, Sung-Suk R (2015) Time series modeling for forecasting land price change rate-focusing on the intervention ARIMA model. Korea Real Estate Acad Rev 60:142–154
17. Wu W (2014) Effectiveness evaluation of housing price regulation policy based on PSR model.business information. Bus Inf 20:63–63
18. Xie X, Hu G (2007) A comparison of Shanghai housing price index forecasting. In: International conference on natural computation, pp 221–225
19. Zhang H, Li Y, Li H (2013) Multi-agent simulation of the dynamic evolutionary process in Chinese urban housing market based on the GIS: the case of Beijing. Autom Constr 35(11):190–198
20. Zhang Z, Tang W (2016) Analysis of spatial patterns of public attention on housing prices in Chinese cities: a web search engine approach. Appl Geogr 70:68–81
21. Zhou Z (2016) Overreaction to policy changes in the housing market: evidence from Shanghai. Reg Sci Urban Econ 58:26–41

A Descriptive Analysis of the Impact of Air Pollution on the Mortality of Urban and Rural Residents in Mianyang

Jianping Xu[1], John Thomas Delaney[2], Xudong Chen[3](✉), and Liming Yao[4]

[1] Regional Economic Development and Enterprises Administration Center, Sichuan University, Chengdu 610062, People's Republic of China
[2] Kogod School of Business, American University, 4400 Massachusetts Avenue, Washington, DC 20016-8044, USA
[3] School of Management Science, Chengdu University of Technology, Chengdu 610062, People's Republic of China
chengxudong198401@163.com
[4] Uncertainty Decision-Making Laboratory, Sichuan University, Chengdu 610064, People's Republic of China

Abstract. The paper uses the air pollution data of Mianyang, a city in Sichuan Province, and the death data of its urban and rural residents from 2008 to 2014 to investigate the relationship between air pollution and human mortality. SPSS19.0 is utilized to conduct a descriptive analysis of the correlation between the principal air pollutants (including PM_{10}, NO_2, SO_2) and the mortality (including gender, age, education, respiratory diseases, non-respiratory diseases, and chronic obstructive pulmonary diseases). Hence, the paper shows to what extent air pollution affects the mortality rate of the people in Mianyang, thereby further evaluating the influence of air pollution on the health of the residents. The results indicate that the rising trend of air pollution is consistent with the expansion of industrial production. The paper estimates that the mortality rate is correlated to air pollution during the normal development of society. The government should emphasize air pollution control and take measures to reduce air pollution, so as to improve the quality of life of urban and rural residents and to reduce the mortality caused by air pollution directly or indirectly.

Keywords: Air pollution · Mortality · Descriptive analysis

1 Introduction

Haze related pollution have been proved to be the reason for lung cancers and some other diseases and also have serious impact on and threatening future health consequences across the country [9]. Numerous Chinese and foreign studies have shown that PM_{10} and gaseous pollutants, such as sulfur dioxide (SO_2) and nitrogen dioxide (NO_2), have a significantly negative impact on human

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_150

health. Moreover, a positive correlation between mortality and air pollution has been demonstrated in multiple researches [1,4,8,16]. The 2002 WHO report estimated that the global urban air pollution caused at least 1 million deaths per year and 7.4 million disability-adjusted life years (DALY) [3].

China is one of the countries with severe environmental problems. The concentration of particulate air pollution is far beyond the developed countries in Europe and North America. In the past 30 years, the economy and industry had been developing continuously. The rapid growth of energy consumption has aggravated air pollution, which becomes one of the major risk factors for damaging human health. The irrationality of the energy structure is that coal accounts for nearly 70% of the total energy consumption in China. The PM and gaseous pollutants emitted from coal combustion are the main sources of air pollution. The cities in China face severe air pollution, and the PM10 is much higher than the 20 $\mu g/m^3$ (annual average) and 50 $\mu g/m^3$ (24 h average)–the air quality standards recommended by the WHO.

Mianyang City, which is located in Southwest China, is the second largest city in Sichuan Province in terms of economic output. It has a dry climate and is densely populated. The agriculture, industry, and service (transportation) are well-developed there with a ratio of 15.3: 50.5: 34.2, indicating that the industry represents half of the economic output. Therefore, it is of significance to examine the impacts of air pollution on human health. In order to study the impact of air pollution on the mortality rate of urban and rural residents in Mianyang, the paper conducts a descriptive analysis, investigating the correlation between the principal air pollution data (including PM_{10}, NO_2, SO_2) and the human mortality.

2 Data and Methods

Mianyang City, the main terrain of which is of hills, is located in the northwest of the Sichuan Basin. It has a subtropical humid monsoon climate. Due to the influence of the northerly air flow, Mianyang is dry and cold, with less rainfall during the winter half year; however, controlled by the southerly air flow in the summer half year, it is hot, rainy, and humid. The household population of Mianyang was 5.488 million by the end of 2015.

The data used in this paper include mortality data and air pollution data from 2008 to 2014.

Mortality data: All deaths from 2008 to 2014 are gathered from the Mianyang municipal government. The paper clarifies the diseases into two groups according to WHO International Classification of Diseases (ICD-10): deaths caused by respiratory diseases (J00∼J98) and deaths caused by other diseases (or non-respiratory diseases). It, then, conducts an individual analysis of the chronic obstructive pulmonary disease (COPD). The human mortality data are classified into the following categories: sex-specific mortality, age-specific mortality, education-specific mortality, and mortality caused by respiratory and non-respiratory diseases (the chronic obstructive pulmonary disease is singled out).

Principal air pollution data: The annual average concentration of PM_{10}, NO_2, SO_2 from 2008 to 2014 are provided by the Mianyang Environmental Protection Bureau.

Methods: SPSS19.0 is utilized to conduct a descriptive analysis of the air pollution data of Mianyang and the mortality of its urban and rural residents from 2008 to 2014.

3 Results

3.1 A Descriptive Analysis of Air Pollution and Death Data

The statistical data of air pollution and death in Mianyang from 2008 to 2014 are shown in Tables 1, 2, and 3. The annual average PM_{10} of Mianyang from 2008 to 2014 was 0.059 mg/m^3, and the peak value was 0.082 mg/m^3 in 2008. The annual average SO_2 concentration was 0.015 mg/m^3, and the peak value was 0.029 mg/m^3 in 2009. The annual average NO_2 concentration was 0.019 mg/m^3, and the peak value was 0.039 mg/m^3 in 2009. The average death toll of men and women in Mianyang was 14,140 and 9061, respectively from 2008 to 2014. The average death toll of people with unknown education was 2098, illiterate or semi-illiterate 7404, primary education 10354, secondary education 2963, university and above 298. The average deaths were 2353 for people aging under 44, 6032 for residents from 45 to 64, and 14818 for those ageing over 65. The average deaths caused by respiratory diseases and non-respiratory diseases were 6663 and 16541, respectively. The average deaths caused by chronic obstructive pulmonary diseases were 767.

It can be seen from Table 1 and the corresponding line graph that the Mianyang air pollution generally showed a declining trend from 2008 to 2014. PM_{10}, SO_2 and NO_2 reached their peak from 2008 to 2009. Then, PM_{10} showed a downward trend till 2014. SO_2 and NO_2 declined moderately from 2009 to 2012. After that, SO_2 showed an upward trend from 2012 to 2014, while NO_2 fluctuated repeatedly between 2012 and 2014.

The paper will discuss, from three aspects, why the overall trend of air pollution in Mianyang declined from 2008 to 2014, even though it fluctuated sometimes during this period.

(1) The peak of air pollution between 2008 and 2009 was caused by the 2008 Wenchuan earthquake. The research shows that due to the earthquake, the collapsed buildings, the leakage of toxic gases and radioactive materials lead to the aggravation of air pollution [2]. Because of the great magnitude and severe damage of the 2008 Wenchuan earthquake, Mianyang, located in Longmenshan earthquake zone, was affected by the aftershocks. This earthquake caused heavy casualties and severe property losses, and Mianyang became one of the hardest hit areas. The urban and rural housing, collapsed or damaged, reached 6,284,011and 58,441,355, respectively. In addition, 47% of the industrial enterprises in Mianyang were affected, with 74% of plant damages. Thus, the collapse

and damage of residential buildings and industrial plants caused by the earthquake, to a certain extent, lead to an increase in airborne dust particles and toxic gas content, and exacerbate air pollution, thereby resulting in the Mianyang air pollution peak between 2008 and 2009.

(2) During the restoration and reconstruction process starting from June, 2008, due to the industrial and agricultural losses and the impacts of the earthquake, the local workers tended to find employment outside Mianyang, whereas the enthusiasm of the non-local job seekers, especially skilled workers and high-tech talents, were seriously affected in a certain period of time. Second, because of the severe capital losses caused by the damages to the plants and the equipment, some enterprises need relocation and reconstruction, increasing the costs and slowing down the process of the industrial and agricultural recovery in Mianyang. Moreover, most of the enterprises need a longer recovery period, since the major coal companies stop production due to the disaster; other kinds of electricity supplies are affected; and the main roads and power facilities are severely damaged [7]. During the recovery period, because of the less industrial and agricultural emissions caused by the decreased productivity of the enterprises, and less traffic pollution emissions, the air pollution generally showed a downward trend from 2009 to 2012. After a 3–4 year of recovery, the overall industrial and agricultural industries were recovered in 2012, and the infrastructure was gradually improved, resulting in a rise in air pollution.

(3) The reasons why the air pollution in Mianyang showed a decreasing trend between 2013 and 2014 are twofold. First, according to Sichuan Air Pollution Prevention and Control, compared to 2012, the annual average concentration of PM_{10} should decrease by 10%. Mianyang, as one of the key cities in Sichuan Province, takes strict emission reduction measures to control the coal-fired elevated-source, industrial point source, urban non-point source, and moving source (vehicle pollution), through strict assessment and accountability. Therefore, air pollution emissions are reduced from a number of sources. Second, because of the overall improvement in the environmental awareness of the Mianyang residents, the development of tourism, the promotion of energy conservation, as well as green travel, air pollution emissions are reduced.

Table 1. Air pollution in Mianyang from 2008 to 2014

Year	PM_{10} (mg/m^3)	SO_2 (mg/m^3)	NO_2 (mg/m^3)
2008	0.082	0.015	0.038
2009	0.081	0.029	0.039
2010	0.05	0.021	0.02
2011	0.041	0.016	0.015
2012	0.051	0.001	0.003
2013	0.071	0.009	0.02
2014	0.035	0.012	0.0009

Table 2. The Death Toll in Mianyang from 2008 to 2014

Year/ death toll	Total population of Mianyang	Male death toll	Female death toll	Unknown degree of education death toll	Illiterate/ semi-literate death	Elementary school education death	Middle school education death	University and above education death
2008	5407100	10360	6504	1867	6912	5245	2569	271
2009	5446500	10557	6698	2047	6781	5620	2625	280
2010	5418700	12129	7553	2543	7648	6194	3025	272
2011	5433600	14387	8999	2706	9449	7490	3453	284
2012	5454000	15395	10044	2980	10433	8197	3520	306
2013	5474000	15942	10382	3243	10605	8458	3725	290
2014	5488000	20212	13252	0	0	31275	1827	386

Table 3. The death toll in Mianyang from 2008 to 2014

Year/ death toll	0–44 years old death toll	45–64 years old death toll	Over 65 year old death toll	Respiratory disease death toll	Non-respiratory disease death toll	COPD (chronic pulmonary disease) death toll
2008	2274	4638	9952	4807	12057	153
2009	2107	4697	10451	4891	12364	143
2010	2415	5392	11874	5377	14305	269
2011	2503	6248	14633	6523	16861	280
2012	2417	6448	16573	7233	18205	519
2013	2343	6671	17309	7700	18623	638
2014	2412	8136	22939	10112	23375	3367

3.2 Description of the Relationship Between Air Pollution and Human Mortality in Mianyang from 2008 to 2014

(Note: In order to better reflect the relationship between air pollution and human mortality, the paper uses air pollution values * % to compare with the mortality rate.)

Figure 1 illustrates that the PM_{10} fluctuated between 2008 and 2014. The peak between 2008 and 2009 is partly because the collapsed buildings, the leakage of toxic gases and radioactive materials, due to the Wenchuan earthquake, aggravated air pollution. PM_{10}, SO_2, and NO_2 rose steadily since 2010, during which the male and female mortality showed the same trend. Specially, the male mortality rate increased greatly, indicating that men are more sensitive to the change in air pollution concentration. Under the impacts of the national policies and local laws and regulations, the air pollution concentration has decreased since 2013. However, there is a greater increase in the mortality rate, which contradicts the conclusion. By searching information and literature, the paper finds that Mianyang faced a severe aging problem from 2012 to 2014. More precisely, Mianyang has a large aged population which grows fast. By the end of 2013, the residents aged 60 and above reached 1,07 million, accounting for 19.8% of the total population, and the individuals who die of old age were 22,939 in 2014. Therefore, the 2013–2014 mortality rate is influenced by the aging population [10,14]. Because of the changes in the social demographic structure, the impact of air pollution on human death is reduced.

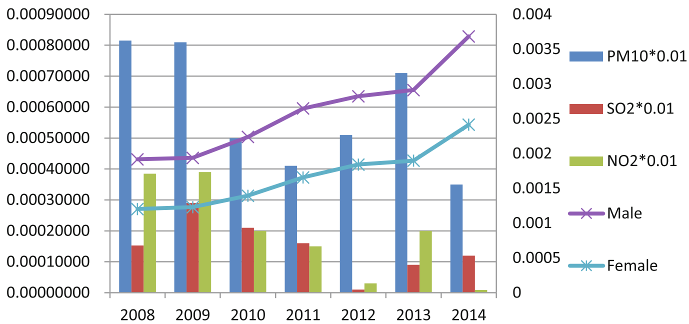


Fig. 1. The relationship between air pollution and gender-specific mortality

Obviously, it can be seen from Fig. 2 that there was a dramatic change in the death rate of residents with primary and secondary education. The reason is because the Mianyang Statistic Bureau changed the classification standard of the education-specific human mortality. That is, people with unknown education, illiterate or semi-illiterate, primary, secondary education, and university and above is changed into junior high and below, technical secondary education, high schools, technical schools, universities and colleges. Hence, between 2013 and 2014, the dramatic changes in human mortality were due to the classification criteria of the Mianyang Statistic Bureau, which was not related to air pollution.

The mortality rate of people with primary school education, illiterate or semi-literate, grew steadily from 2008 to 2013, whereas that of residents with secondary education, university and above showed a stable trend, maintaining the same level. It indicates those in terms of the same air pollution level, the higher the education level, the stronger the awareness of environmental protection. Therefore, these residents tend to take preventive measures more timely and effectively when exposed to air pollution, and they are less affected [5,11]. The impact of air pollution on human death is positively correlated with the educational level of the residents. The more serious the air pollution is, the higher the mortality rate is. However, due to the differences in people's education level, the awareness of air pollution prevention varies considerably. The people with the lower education level may have the lower residents' awareness of air pollution prevention, which results in a higher chance of being affected by the air pollution.

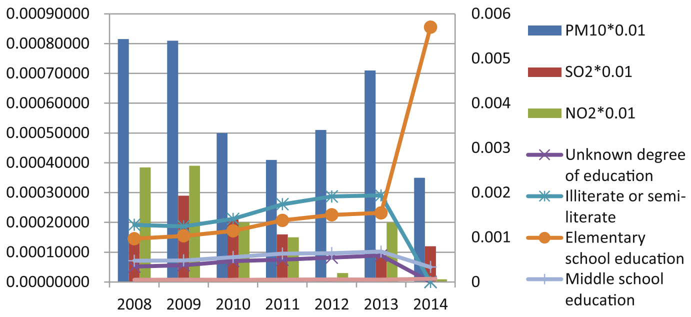


Fig. 2. The relationship between air pollution and education-specific mortality

According to Fig. 3, between 2008 and 2014, the mortality rate of residents aged 0–44 years and 45–64 years showed a stable trend, while that of people aged 65 years and over was consistent with the trend of air pollution between 2008 and 2013, indicating that air pollution impacted individuals aged 65 years and above the most. In terms of residents aged 44 and over, their physical function declines with age, and they are more easily affected by the external environment. Air pollution, to a larger extent, impacts individuals aged 44 and over, the death rate of whom may increase with a rise of the air pollution concentration [13]. However, the mortality rate of people aged 65 years and over in 2013–2014 showed a reverse trend with air pollution, which contradicts the argument that the senior residents are susceptible to air pollution. By searching literature, the paper finds that Mianyang faced a severe aging problem from 2012 to 2014. By the end of 2013, the residents aged 60 and above reached 1,07 million, accounting for 19.8% of the total population. Hence, a high aging population impacts the human mortality.

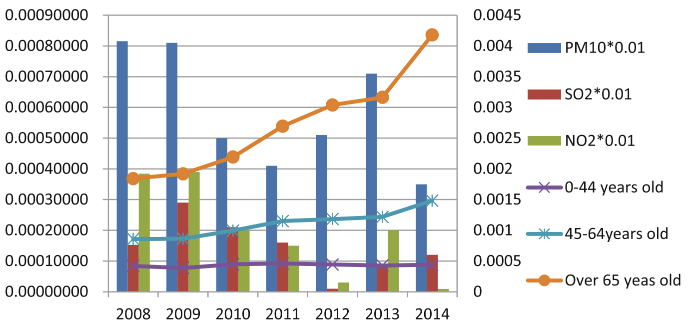


Fig. 3. The relationship between air and age-specific human mortality

It can be seen from Fig. 4 that the mortality rate caused by respiratory diseases showed a steady upward trend from 2008 to 2013, but it grew rapidly from 2013 to 2014. Due to the 2008 Wenchuan earthquake, Mianyang experienced a recovery period of agriculture, industry, and transport infrastructure. As a result, air pollution declined from 2009 to 2011. Nevertheless, the concentration of air pollution increased from 2011 to 2013, which correspondingly saw a rising trend of respiratory diseases-related deaths. It indicates that air pollution has a significant impact on the residents who died of respiratory diseases. During the period of 2013 and 2014, the human mortality rate caused by respiratory diseases showed a reverse trend with air pollution, contradicting the argument that there is, according to existing researches, a positive correlation between respiratory diseases-related deaths and air pollution [6,12,15]. From the Mianyang 2013–2014 government and other reports, the paper finds that this city faced a severe aging problem. Due to the prevalence of respiratory diseases among the senior residents, the 2013–2014 death rate was correlated considerably with the aging population, less with air pollution.

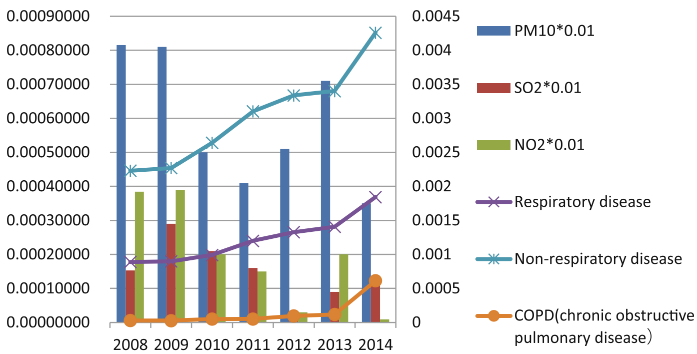


Fig. 4. The relationship between air and mortality of disease

4 Discussion

The paper uses descriptive statistical analysis to investigate the correlation between major air pollutants and residents' mortality in Mianyang City. Given the impacts of natural disasters, national policies, and socio-demographic changes, the paper concludes that the urban and rural residents, who are male, illiterate or semi-literate, or with primary education, the middle-aged or aged, or with respiratory diseases, are more susceptible to the increase in the PM_{10}, NO_2 and concentration.

The accumulation of pollutants and photochemical reactions in various cities leads to complex regional air pollution issues. Mianyang, located in the northwest of Sichuan Basin, has a subtropical humid monsoon climate. The hot, rainy, and wet weather is not conducive to the diffusion of particulate matters, aggravating the air pollution of Mianyang. Furthermore, factors, such a high level exploitation of the rich mineral resources in Mianyang and its emphasis on industry development, contribute to the severe air pollution. Therefore, the Mianyang government should consider the health risk of air pollution, and take effective and feasible measures to tackle this issue, thereby reducing the pollution-related health risks of its urban and rural residents.

Here are four policy suggestions we want Mianyang government can accept and conscientiously implement. First, government needs to improve the local industrial structure by circulation type industrial building, which can reduce air pollution and improve resource utilization efficiency, also can achieve the transformation from primary industry into tertiary industry. Second, attaching great importance to the city public transport priority, vigorously implementing the bus priority development strategy, and increasing the city public traffic investment and construction strength in order to reduce harmful automobile exhaust emissions and protect residents' health. Third, promotion of industrialized and enterprise and market-oriented pollution treatment, encouragement of fund raised from diverse channels and various quarters in society, and intensified collection of charges on urban sewage treatment and garbage disposal, which can effectively reduce every kind of contaminative source. Last but not least, Mianyang government should make the enactment of relevant laws and regulation and raise residents' awareness about the health risks of air pollution to reduce the risk of death for the people, who are male, illiterate or semi-literate, or with primary education, the middle-aged or aged, or with respiratory diseases.

In addition, the study has certain limitations: First of all, due to the high mobility of the Mianyang urban and rural household population, the reported death information may be false or incorrect. Meanwhile, there may be a certain degree of misclassification of death data, which might affect the accuracy of residents' mortality. Secondly, the air pollution concentration measured at the fixed sites represents the exposure level of the urban and rural residents, which may differ, to some extent, from the real exposure level. Moreover, due to damage and impact of the Wenchuan earthquake, there might be a certain deviation of air pollution statistical data gathered at each site, affecting the analysis results.

Acknowledgements. This research was supported by Sichuan Social Science Planning Program Key Project (Grant No. SC16A006), System Science and Enterprise Development Research Center Project (Grant No. Xq16C12) and Sichuan University Central University Research Fund project (Grant No. skqy201653), National Natural Science Foundation for Young Scholars of China (Grant No. 71301109), Soft Science Program of Sichuan Province (Grant No. 2017ZR0154), and Sichuan University (Grant No. skqx201726).

References

1. Fischer PH, Marra M et al (2015) Air pollution and mortality in seven million adults: the Dutch environmental longitudinal study (duels). Environ Health Perspect 123(7):697–704
2. Gotoh T, Nishimura T et al (2002) Air pollution by concrete dust from the Great Hanshin Earthquake. J Environ Qual 31(31):718–723

3. Gulland A (2002) Air pollution responsible for 600 000 premature deaths worldwide. BMJ (Clinical Res Ed) 325(7377):1380
4. Jerrett M, Burnett RT et al (2013) Spatial analysis of air pollution and mortality in Los Angeles. Am J Respir Critical Care Med 188(5):727–736
5. Kan H, Chen B (2008) Season, sex, age, and education as modifiers of the effects of outdoor air pollution on daily mortality in Shanghai, China: The public health and air pollution in Asia (PAPA) study. Environ Health Persp 116(9):1183–1188
6. Katanoda K, Sobue T, Satoh H (2011) An association between long-term exposure to ambient air pollution and mortality from lung cancer and respiratory diseases in Japan. J Epidemiol 21(2):132–143
7. Kirby E (2008) Stress changes from the 2008 Wenchuan earthquake and increased hazard in the Sichuan basin. Nature 454(7203):509–510
8. Lelieveld J, Evans JS et al (2015) The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525(7569):367–371
9. Li Y, Xu Y et al (2016) Haze-related air pollution and the tourism industry in China. In: GAI international academic conferences proceedings, pp 122–131
10. Markides KS, Eschbach K (2005) Aging, migration, and mortality: current status of research on the Hispanic paradox. J Gerontol 60(Spec No 2):68
11. O'Neill MS, Bell ML et al (2008) Air pollution and mortality in Latin America: the role of education. Epidemiology 19(6):810–819
12. Saldiva PH, Lichtenfels AJ et al (1994) Association between air pollution and mortality due to respiratory diseases in children in São Paulo, Brazil: a preliminary report. Environ Res 65(2):218–225
13. Schwartz J, Dockery DW (1992) Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145(3):600–604
14. Strehler BL, Mildvan AS (1960) General theory of mortality and aging. Science 132(3418):14
15. Wong TW, Tam WS et al (2002) Associations between daily mortalities from respiratory and cardiovascular diseases and air pollution in Hong Kong, China. Occup Environ Med 59(1):30–35
16. Xie W, Li G et al (2015) Relationship between fine particulate air pollution and ischaemic heart disease morbidity and mortality. Heart 101(4):257–263

Author Index

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0